DOCUMENT RESUME

ED 404 371 TM 026 457

AUTHOR Muthen, Bengt; And Others
TITLE Opportunity-to-Learn Effects on Achievement: Analytical Aspects.
INSTITUTION National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
PUB DATE Jul 95
CONTRACT RS90159001
NOTE 53p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; *Achievement Tests; Educational Assessment; *Mathematics Tests; National Surveys; *Performance Factors; Scores; Secondary Education; Structural Equation Models; Tables (Data); Test Construction; *Test Items
IDENTIFIERS Large Scale Programs; *National Assessment of Educational Progress; National Education Longitudinal Study 1988; *Opportunity to Learn

ABSTRACT

A set of methods is proposed for the analysis of opportunity to learn (OTL) in relation to achievement in large-scale educational assessments. The focus is on how to assess the effect of OTL on performance while taking prior performance and other background factors into account. The methods are illustrated with mathematics data from the National Assessment of Educational Progress and from the National Education Longitudinal Study. Methods are discussed for combining OTL information, for studying the OTL sensitivity of test items, and for studying OTL effects in the context of multivariate proficiency scores as well as scores from several occasions. Implications for future large-scale educational assessments are discussed. An appendix presents input for structural equation modeling through the LISCOMP computer program. (Contains 5 figures, 13 tables, and 16 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TM

CRESST

TM

ED 404 371

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Opportunity-to-Learn Effects on Achievement: Analytical Aspects

Bengt Muthén, Li-Chiao Huang, Booil Jo, Siek-Toon Khoo, Ginger Nelson Goff, John Novak, and Jeff Shih

CRESST/University of California, Los Angeles
Graduate School of Education & Information Studies

Bengt Muthén, Project Director

TM026457

- UCLA Center for the Study of Evaluation

In collaboration with:

- University of Colorado
- NORC, University of Chicago
- LRDC, University of Pittsburgh
- University of California, Santa Barbara
- University of Southern California
- The RAND Corporation

BEST COPY AVAILABLE

National Center for Research on
Evaluation, Standards, and Student Testing

Technical Review Panel for Assessing the Validity of
the National Assessment of Educational Progress

Opportunity-to-Learn Effects on Achievement: Analytical Aspects

Bengt Muthén, Li-Chiao Huang, Booil Jo, Siek-Toon Khoo,
Ginger Nelson Goff, John Novak, and Jeff Shih

CRESST/University of California, Los Angeles
Graduate School of Education & Information Studies

Bengt Muthén, Project Director

Final Deliverable

July 1995

U.S. Department of Education
National Center for Education Statistics
Grant RS90159001

Center for the Study of Evaluation
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1521
(310) 206-1532

The research reported herein was supported under the National Center for Education Statistics Contract No. RS90159001 as administered by the Office of Educational Research and Improvement, U.S. Department of Education. The findings and opinions expressed in this report do not reflect the position or policy of the National Center for Education Statistics, the Office of Educational Research and Improvement or the U.S. Department of Education.

Abstract

A set of methods is proposed for the analysis of opportunity to learn (OTL) in relation to achievement in large-scale educational assessments. The focus is on how to assess the effect of OTL on performance while taking prior performance and other background factors into account. The methods are illustrated with mathematics data from the National Assessment of Educational Progress and from the National Education Longitudinal Study. Methods are discussed for combining OTL information, for studying the OTL sensitivity of test items, and for studying OTL effects in the context of multivariate proficiency scores as well as scores from several occasions. Implications for future large-scale educational assessments are discussed.

OPPORTUNITY-TO-LEARN EFFECTS ON ACHIEVEMENT: ANALYTICAL ASPECTS[1]

Bengt Muthén, Li-Chiao Huang, Booil Jo, Siek-Toon Khoo, Ginger Nelson Goff, John Novak, and Jeff Shih

CRESST/UCLA Graduate School of Education & Information Studies

Introduction

Interest in opportunity-to-learn (OTL) standards is continuing to grow, in part because of new legislation related to Goals 2000 and the Elementary and Secondary Education Act. OTL issues are embedded in a large set of interrelated policy, conceptual, and measurement issues. For a review, see McDonnell (in press). In our article we consider OTL in the context of performance analyses in large-scale assessments. We are interested in OTL in terms of school- and classroom-related activities including content exposure and coverage as well as instructional conditions, practices, and processes. The focus is on analytical issues that arise when

> changing the question from "What students know and can do" to "What students know and can do as a result of their educational experiences." (Burstein & Winters, 1994)

The quotation can be translated into terms of regression analysis. The assessment of OTL effects can be described by three key constructs: Performance is the outcome and OTL and prior performance are the predictors. OTL and prior performance are usually correlated. We should therefore control for prior performance and focus on the partial effect of OTL. This analytical task is not, however, straightforward.

As an example, consider results from the National Assessment of Educational Progress (NAEP). These are typically reported as distributions for various subgroups of the population defined by gender, ethnicity, parental

[1] We are thankful to Leigh Burstein for many helpful comments on our initial work. We also thank Lynn Orlando for her assistance.

education, type of community, type of school, region of the country, and so forth. Take as a specific example comparisons of ethnicity subgroup performance on algebra among 8th graders in the 1992 NAEP main assessment. It is well known that the performance differences are large among these subgroups when other factors are not taken into account. One may ask to which extent the performance differences are reduced when taking across-group variation in OTL into account.

The across-group variation is quite large for a key OTL variable: 37% of Black and 33% of Hispanic students report that they are in prealgebra or algebra classes, whereas 53% of Whites and 63% of Asians report that they are in such 8th-grade classes. Should NAEP report OTL-conditioned results in addition to the usual unconditioned results? For a discussion of some issues of comparing performance when attempting to control for curriculum, see the Baker (1993) and Westbury (1993) debate in the context of comparing American and Japanese 8th-grade math performance. Figure 1 shows what happens in the context of NAEP algebra performance when controlling for algebra OTL. Box-and-whiskers plots show the subgroup distributions both overall (left part of the figure) and for students who are in prealgebra or algebra classes (right part of the figure). The figure shows that subgroup differences are not reduced but remain almost the same when conditioning on OTL. All four subgroups have about the same increase in algebra performance when changing the study from all students to students in prealgebra or algebra classes. This result does not change when one excludes the prealgebra group and focuses on the algebra group.

One may consider several hypotheses for this finding. First, the prealgebra/algebra students may not have the same 7th-grade math ability across subgroups, or may be otherwise different in their preparation prior to 8th-grade algebra studies. This means that although the OTL effects are the same across subgroups, the starting points are different. The mechanisms for selection into prealgebra/algebra classes would then be different across subgroups; see also Kifer (1992). Second, the quality of OTL may differ across subgroups. For example, the starting points may be the same, but the OTL quality may differ across subgroups. Third, the OTL reporting may differ across subgroups in terms of reliability and validity. This example illustrates the difficulty in analyzing how performance relates to OTL. Prior performance and other relevant background factors need to be controlled for and OTL needs to be measured in detail and with precision. This can provide for interesting analyses of the various hypotheses.

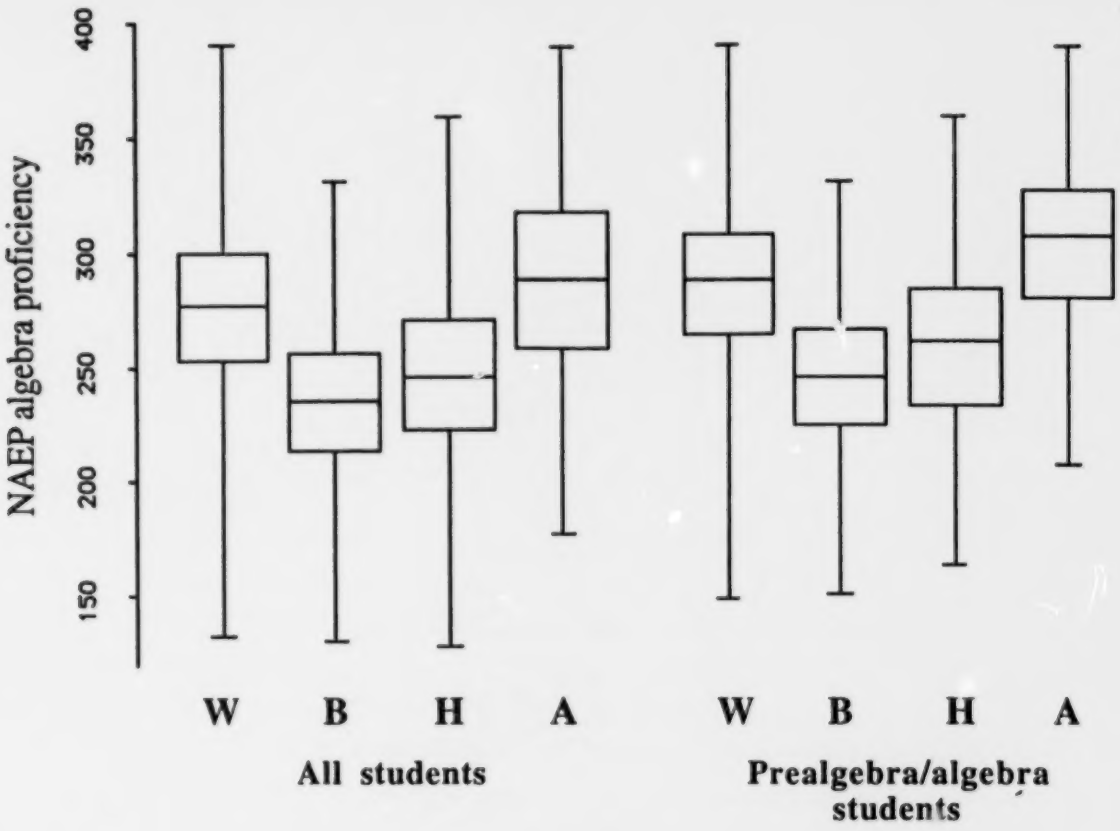


Figure 1. NAEP 1992 8th-grade algebra proficiency related to ethnicity (White, Black, Hispanic, Asian) and algebra class type.

This article presents some ideas for studying OTL effects using latent variable modeling. An important theme is the notion of a general factor influencing achievement. A general factor notion is introduced for two reasons. First, OTL effects can be clearly described as effects that go beyond what is expected by the general factor. Such added effects can be described in terms of outcome variables corresponding to specific factors—variables affecting more narrow components of the test performance. Second, it is more likely that OTL effects can be found with respect to specific factors as opposed to a general factor (e.g., an overall math score).

Controlling for initial ability is possible with longitudinal achievement data such as the National Education Longitudinal Study of 1988 (NELS) or the

Longitudinal Study of American Youth (LSAY), but longitudinal data are seldom available in performance assessments. An interesting question is what one can do given data from a cross-sectional assessment such as NAEP. This article considers to which extent the general factor is a good proxy for prior performance.

To assess the effects of OTL as well as possible, this article considers analyses with multivariate information on OTL, performance, and prior performance, as well as information from multiple occasions. Not only can OTL effects be studied with the help of such multivariate modeling, but scores adjusted for OTL or lack thereof can also be derived for reporting of assessments. With better OTL measures in future large-scale assessments, the OTL effects on the specific factors can provide useful indicators of the performance effects of population changes in OTL.

The article reports on analytic work with the 1992 NAEP and the 1988-1990 NELS mathematics data, divided into analyses of OTL, analyses of performance, and relating the two constructs by multivariate modeling.

Analyses of Multivariate OTL Information

Analysis Goals

As a first step in studying OTL effects on mathematics achievement, the OTL information itself needs to be carefully analyzed. In line with Burstein (personal communication, April 1994), we argue that multivariate information should be used to more fully characterize types of math classes rather than using isolated pieces of information based on single questionnaire items. To illustrate this, we present analyses based on student and teacher reports for both NAEP and NELS data.

In the Second International Mathematics Study (SIMS; see, e.g., Burstein, 1992), four 8th-grade class types—remedial, typical, enriched, and algebra—were created from teacher questionnaire data and information on textbooks used. These class-type variables provided valuable information in the SIMS achievement analyses. A corresponding classification is not, however, readily available based on NAEP and NELS data. We will attempt to derive a class-type variable similar to the one used in SIMS using teacher-reported information on math emphasis variables in NAEP and NELS.

Methods

Data from the National Assessment of Educational Progress (NAEP) are obtained as a multistage, national probability sample for Grades 4, 8, and 12. The 1992 main assessment data that will be used here cover five content areas in mathematics: numbers and operations (arithmetic), measurement, geometry, data analysis and statistics, and algebra. Test results were obtained for almost 10,000 students per grade. The analyses in this article will focus on 8th-grade and 12th-grade data. The National Education Longitudinal Study (NELS) is another nationally representative achievement study. In it, over 20,000 students are tested. The survey was first administered in 1988 for 8th graders with follow-up tests in the 10th and 12th grades.

NAEP 1992 8th-grade math data contain student-reported information that classifies the student as belonging to a class where no or regular 8th-grade math is taught (50%), where prealgebra is taught (27%), or where algebra is taught (23%). In addition, there are teacher-reported math emphasis variables corresponding to the five content areas of numbers and operations, measurement, geometry, data analysis and statistics, and algebra, as well as emphasis on learning skills/procedures, reasoning/analysis, communication, appreciating math, and teacher-determined class ability level. NELS data contain similar information. These variables have the potential of giving more information on class type than the student-administered question. Factor analyses are carried out using maximum-likelihood estimation and oblique rotations using the promax method.

Results

The teacher-reported variables for NAEP are described in Table 1. As shown in Table 1, factor analysis of the teacher-reported variables in NAEP indicates a clear four-factor solution with one factor corresponding to emphasis on reasoning, communication, and appreciating math (a factor related to NCTM goals), and three other factors corresponding to membership in remedial or typical classes, enriched classes, or algebra classes. The factor loading pattern agrees with usual notions: in remedial and typical classes there is an emphasis on numbers and operations (arithmetic) and facts and skills; in enriched classes there is an emphasis on measurement, geometry, and data analysis and statistics; and in algebra classes there is an emphasis on algebra and reasoning (here the teacher typically also assesses the class ability level as high).

Table 1
NAEP 1992 Grade 8 Factor Analysis of Teacher-Reported Math Emphases

Variables	Factor			
	NCTM	Remedial & Typical	Enriched	Algebra
How much emphasis on:[a]				
Numbers and Operations	-0.02	0.60	0.06	-0.09
Measurement	0.17	0.13	-0.53	-0.23
Geometry	0.00	0.01	0.71	-0.03
Data/Stats/Probability	0.14	-0.04	0.53	0.16
Algebra/Functions	0.11	-0.03	0.03	0.66
Learning facts/Concepts	0.12	0.68	-0.02	-0.05
Learning skills/Procedures	0.21	0.56	0.02	0.04
Reasoning/Analysis	0.48	0.12	0.21	0.45
Communicating math ideas	0.73	0.06	0.11	0.15
Appreciating math	0.67	0.22	0.12	0.11
Math ability of students in class[b]	0.13	-0.11	-0.10	0.59

[a] Response coding: 1. Little or no emphasis. 2. Moderate emphasis. 3. Heavy emphasis.

[b] Math ability coding: 1. Mostly low ability. 2. Mixed ability levels. 3. Most average ability. 4. Mostly high ability.

The Table 1 factor solution can be used to classify students into three class types: remedial and typical, enriched, and algebra. We do so by using standardized factor score values for each of these three factors to classify a student into a class type for which his or her value is the largest. Table 2 shows that this classification appears to have at least a minimal amount of validity in the sense that performance on geometry, measurement, and data analysis and statistics is higher for the enriched group than for the remedial and typical group, and performance on algebra is highest for the algebra group.

The teacher-based classification can be compared to that of the student-based class membership with the three categories: no math this year or 8th-grade math; prealgebra; algebra. Table 3 shows that the two classification schemes give partially different results, calling into question the reliability and validity of the information. It is worthwhile to explore both schemes in relation to performance.

Table 2

Performance Means and Standard Deviations for Class Types Derived From Teacher-Based Emphasis Information

Teacher-reported class type	Content				
	Num & Ops	Measure-ment	Geometry	Data & Stat	Algebra
Remedial & Typical					
Mean	264.05	255.21	255.48	260.03	258.19
SD	29.59	38.04	28.81	35.51	30.51
Enriched					
Mean	268.87	261.94	259.36	263.73	262.05
SD	30.57	38.84	29.77	36.04	31.55
Algebra					
Mean	293.05	291.86	283.14	294.19	292.11
SD	29.35	38.21	29.28	34.94	32.11

Table 3

NAEP 1992 Student-Reported by Teacher-Reported Math Class Type

Student-reported class type	Teacher-reported class type			
	Remedial & Typical	Enriched	Algebra	Total
Grade 8 math / no math				
Number of subjects	1,292	845	311	2,448
% of total				49.77
Prealgebra				
Number of subjects	426	464	452	1,342
% of total				27.28
Algebra				
Number of subjects	115	144	870	1,129
% of total				22.95
Total				
Number of subjects	1,833	1,453	1,633	4,919
% of total	37.26	29.54	33.20	100.00

Turning to NELS88 8th-grade teacher-reported data, Table 4 suggests a similar pattern of emphases as for the NAEP data: In remedial classes common and decimal fractions are emphasized; in typical classes ratio, proportion, and percent problems are emphasized; enriched classes emphasize measurement, geometry, probability and statistics and are characterized by high textbook coverage; algebra classes emphasize algebra and integer problems. Using a similar factor-score based classification as for NAEP, we obtain an almost similar percentage distribution over the class types remedial and typical (44%), enriched (25%), and algebra (31%).

It is clear that the information about 8th-grade class type is not as precise as would be desirable in NAEP and NELS. Given the well-known problem of attenuation in regression estimates resulting from unreliability of predictors, this imprecision will make it harder to show OTL effects on performance. Much more detailed information is also needed. The analytical approaches used here can, however, serve as suggestions for methodologies that can be applied to future data

Table 4
NELS '88 Grade 8 Factor Analysis of Teacher-Reported Math Emphases

Variables	Factor			
	Remedial	Typical	Enriched	Algebra
Emphasis given to:[a]				
Common Fractions	0.92	0.03	-0.02	-0.02
Decimal Fractions	0.92	0.01	-0.02	0.05
Ratio and Proportion	0.13	0.59	0.09	0.13
Percent	-0.02	1.02	-0.03	-0.07
Measurement	0.21	0.02	0.58	-0.17
Geometry	-0.03	0.11	0.64	0.06
Algebra	-0.08	-0.05	-0.01	0.67
Integers	0.09	0.09	0.07	0.66
Probability / Statistics	-0.03	-0.03	0.62	0.18
Problem Solving	0.02	-0.03	0.18	0.21
Percent of textbook coverage[b]	-0.10	-0.01	0.30	0.20

[a] Response coding: 0. Not covered. 1. Minor topic or review topic only. 2. Major topic.

[b] Textbook coverage coding: a. 0–49%. b. 50–59%. c. 60–69%. d. 70–79%. e. 80–89%. f. 90–99%. g. 100%.

BEST COPY AVAILABLE

and that will be more likely to show interesting effects the better the OTL measures.

Analyses of Test Items: OTL-Sensitive items

Analysis Goals

For assessments such as NAEP and NELS, differential item functioning (DIF) is typically investigated for every item, answering the question "Do the items function differently for different gender, ethnicity, and so forth?" If they do, items are discarded. This investigation is not customarily done with respect to groupings based on OTL categories. We suggest, however, that such an analysis is of great interest when the focus is on understanding OTL effects. For example, 8th-grade algebra class membership should make most algebra items easier. This advantage is, however, confounded with the higher math ability of students in algebra classes that customarily results from tracking and other selection mechanisms. DIF analysis takes into account such group differences and can describe item performance differences resulting from OTL, properly conditioned on ability. What DIF analysis searches for are algebra items that are particularly sensitive (or insensitive) to algebra instruction, so that for these items algebra students have an advantage (or disadvantage) that is larger than that on the overall math test.

It is of interest to search for such OTL-DIF items and characterize them. This can be helpful in terms of test construction. OTL-sensitive items may be desirable or undesirable depending on the purpose of the test. If there are few OTL-sensitive items in a test, the test makes for a fair comparison of students who differ in their OTL. However, with few OTL-sensitive items, the test is an insensitive indicator to change in OTL.

Further issues arise related to OTL-sensitive items. Instead of discarding items showing DIF, one could entertain the provocative idea of using adjusted scores by allowing these items to have different difficulty parameters. We are then attempting to measure "potential": What can these students do given opportunities to learn? One could argue that with persons getting such items right, the ones not in algebra classes should get higher scores than those in algebra classes. And, if students not in algebra classes do not get all such items right, their scores should not be as low as students in algebra classes with the same responses. This information is useful in addition to knowing the actual proficiency.

Methods

OTL-DIF analyses can be carried out for the NAEP 8th-grade math items with respect to the three algebra class type categories reported by students (no or 8th-grade math, prealgebra, and algebra) and with respect to membership in enriched classes using the teacher-based composites discussed above. Such analyses can be carried out in several ways. A visually instructive way is to plot each item's proportion correct for the two groups (see, e.g., Bejar, 1980). We present the proportions here in a logit scale to improve linearity. A line can be placed through the scatter of item values and items deviating from the line show especially strong advantage (or disadvantage) in performance by one of the groups over the other. This may indicate deviations from invariance of measurement characteristics (i.e., DIF) and possibly lack of unidimensionality. It should be noted, however, that it is well known that this intuitively useful method for identifying items with DIF is not optimal for tests with items showing large amounts of guessing or strong variations in the item discrimination values. To take such features into account, it is better to carry out Item Response Theory (IRT) analyses with three-parameter logistic models applied to the two groups. As an alternative, the Mantel-Haenszel DIF method, which is the standard DIF method used by ETS for NAEP, may be used. As is the practice at ETS, in this article each item is analyzed as it appears in six different blocks, using the block score as matching variable. For a discussion of DIF methods, see Holland and Wainer (1993).

Results

Figure 2 shows a plot of the 28 multiple-choice and short constructed-response algebra items in the 1992 NAEP main math test. Proportion correct is computed using the "grade only" sample. It is important to note that the line is fitted using all 183 8th-grade math items, so that the algebra item performance is related to the overall math performance. Here, the group of algebra students is compared to the nonalgebra group, excluding the prealgebra students. Items would lie along the broken line if nonalgebra students on the whole performed as well on the math test as the algebra students. Compared to the broken line, the solid line shows that the algebra group performs better overall. The nonalgebra group has an average p value of 0.52 on the 183 items, while the algebra group has an average p value of 0.71. What is of most interest is that the majority of the items are above the solid line and only about a fourth are below it. The items

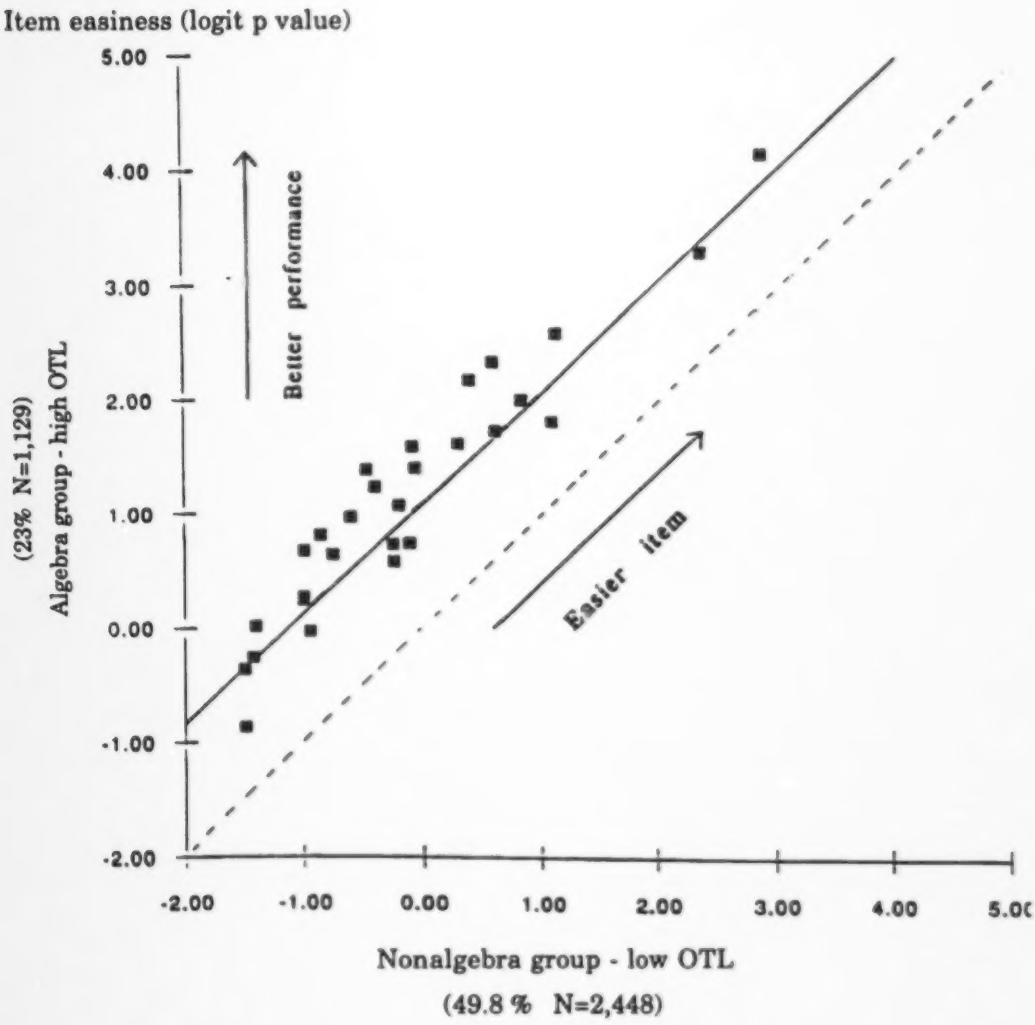


Figure 2. NAEP 1992 8th-grade algebra item performance related to algebra and nonalgebra class type (plot of 28 algebra items).

furthest above the line are especially OTL-sensitive in that performance on these items is more enhanced by belonging to an algebra class than performance on other algebra items. The items below the line indicate that performance on them is enhanced by belonging to an algebra class in that the points are above the broken line, but not enhanced as much as the overall performance advantage for algebra classes would predict.

In Figure 3 the corresponding plot is shown for the prealgebra group compared to the nonalgebra group. In comparing Figures 2 and 3, it is clear that prealgebra has considerably less effect than algebra in enhancing algebra item performance. To get a more detailed assessment of OTL sensitivity in the algebra items, a three-parameter IRT analysis was also carried out. An advantage of such an analysis is that it is possible to take into account group differences in the guessing and discrimination parameters. It may, for example, be the case that on the whole, the nonalgebra group has lower slopes than the algebra group because the topics are less familiar to the nonalgebra group and therefore elicits more measurement error, or random responses.

Table 5 gives the 28 algebra items. The estimated item parameters from each group were linked to a common scale using an analysis of all 183 math items. This takes into account that the two groups differ in their math achievement mean and variance. Although, the algebra group has a higher mean than the nonalgebra group, one might expect that for a given math achievement level, the algebra students would have an advantage on some or most of the algebra items. This would result in different item characteristic curves. Figure 4 shows the item characteristic curves. The remarkable finding is that the curves are almost the same across the algebra and nonalgebra groups. Chi-square testing of invariance of the three parameters across the two groups shows only one significant item (item 18). This says that contrary to expectation, for given achievement level, algebra students do not have a higher probability of giving a correct answer to any of these algebra items. Beyond the higher math achievement mean for algebra students, there is no added advantage of algebra class membership for performance on these algebra items and the higher achievement mean may be at most a result of selection effects. There is therefore no clear evidence of specific algebra learning for algebra students. This may be because the algebra items that were chosen for the 1992 NAEP were quite general and did not require algebra-specific training. These algebra items may therefore be viewed as "OTL-insensitive" items. On the other hand, the grouping of students into algebra and nonalgebra class types may give too coarse of an indicator of algebra OTL given that there may be a great deal of variation in the actual OTL that a student receives.

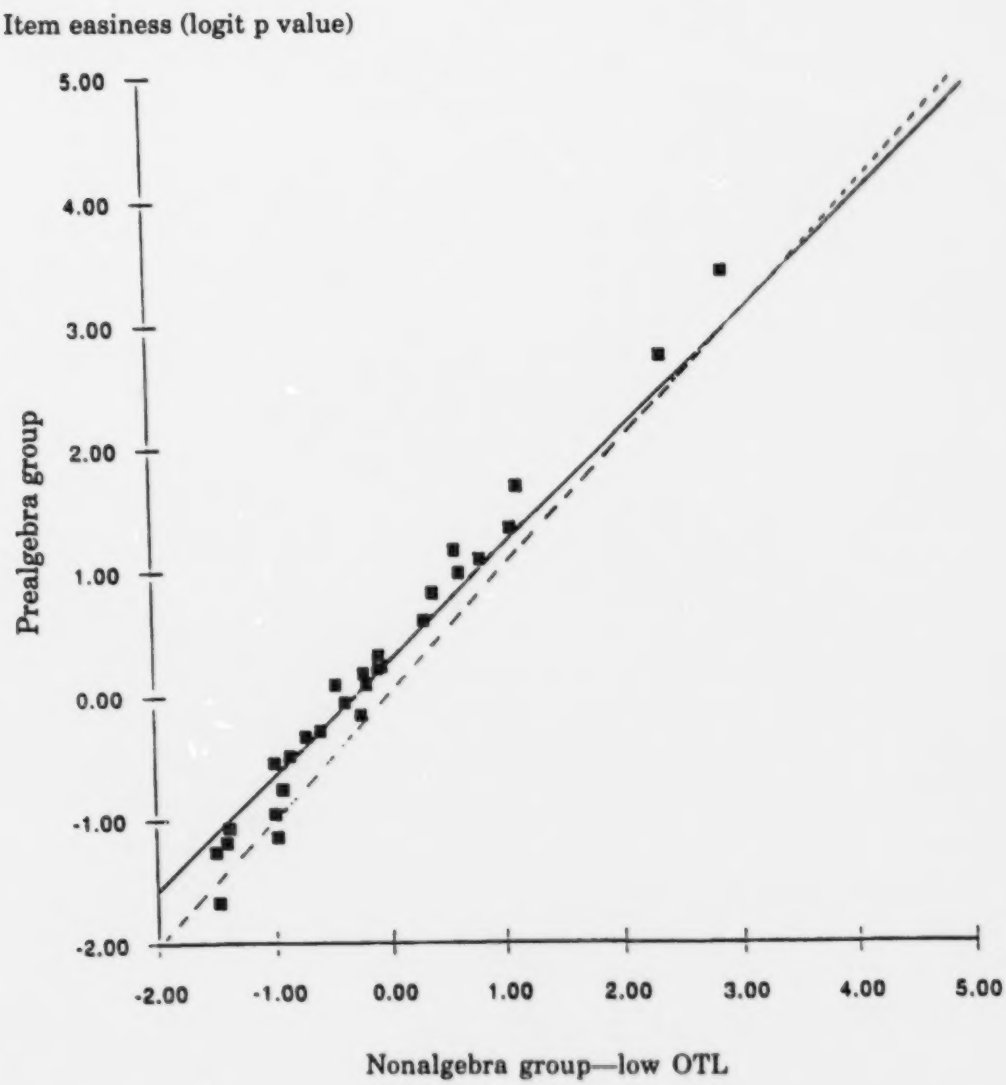


Figure 3. NAEP 1992 8th-grade algebra item performance related to prealgebra and nonalgebra class type (plot of 28 algebra items).

Table 5

28 8th-Grade Algebra Items

Item	NAEPid	Block	Item	
1	M050601C	3	8	FIND (X,Y) SOLUTION—LIN. EQ.
2	M050701C	3	9	TRANSLATE WORDS TO SYMBOLS
3	M050801C	3	10	FIND NUMBER DIAGONALS—POLYGON
4	M018301D	4	10	APPLY CONCEPT OF EQUALITY
5	M018701D	4	14	SOLVE AN INEQUALITY
6	M018801D	4	15	IDENTIFY COORDINATES ON A GRID
7	M019301D	4	20	FIT EQUATION TO DATA
8	M022101E	5	3	TO CONTINUE THE PATTERN, FIGURE B WOULD
9	M022401E	5	6	TOTAL # NEWSPAPERS LEE DELIVERS IN 5 DAYS
10	M023201E	5	15	IF PATTERN CONTINUES, PUPPY WILL WEIGH
11	M019701F	6	1	SOLVE A NUMBER SENTENCE
12	M020401F	6	8	COMPLETE A LETTER PATTERN
13	M021201F	6	14	GRAPH AN INEQUALITY
14	M045201G	7	8	APPLY PATTERN RECOGNITION
15	M045701G	7	11	SOLVE FOR VALUES THAT MAKE AN INEQUALITY
16	M012231H	8	1	USE ORDER OF OPERATIONS
17	M013231H	8	11	EXTRAPOLATE NUMER PATTERN
18	M013731H	8	18	CONVERT TEMPERATURES
19	M052501I	9	3	SELECT GRAPH FOR INEQUALITY
20	M047601K	11	15	SOLVE EQUATION FOR A VARIABLE
21	M047701K	11	16	EXTEND A PATTERN AND COMPUTE
22	M053501L	12	1	UNDERSTAND CONCEPT OF VARIABLE
23	M051701M	13	6	SOLVE EQUATION WITH SQUARE ROOT
24	M052101M	13	10	LOCATE OBJECT ON A GRID
25	M055001N	14	4	EVALUATE EXPRESSION USING ORDER OF OPER.
26	M048801O	15	8	SELECT REASONABLE UNIT OF MEASURE
27	M049401O	15	13	SHOW UNDERSTANDING OF CIRCLE
28	M049601O	15	15	COMPLETE PATTERN IN A TABLE

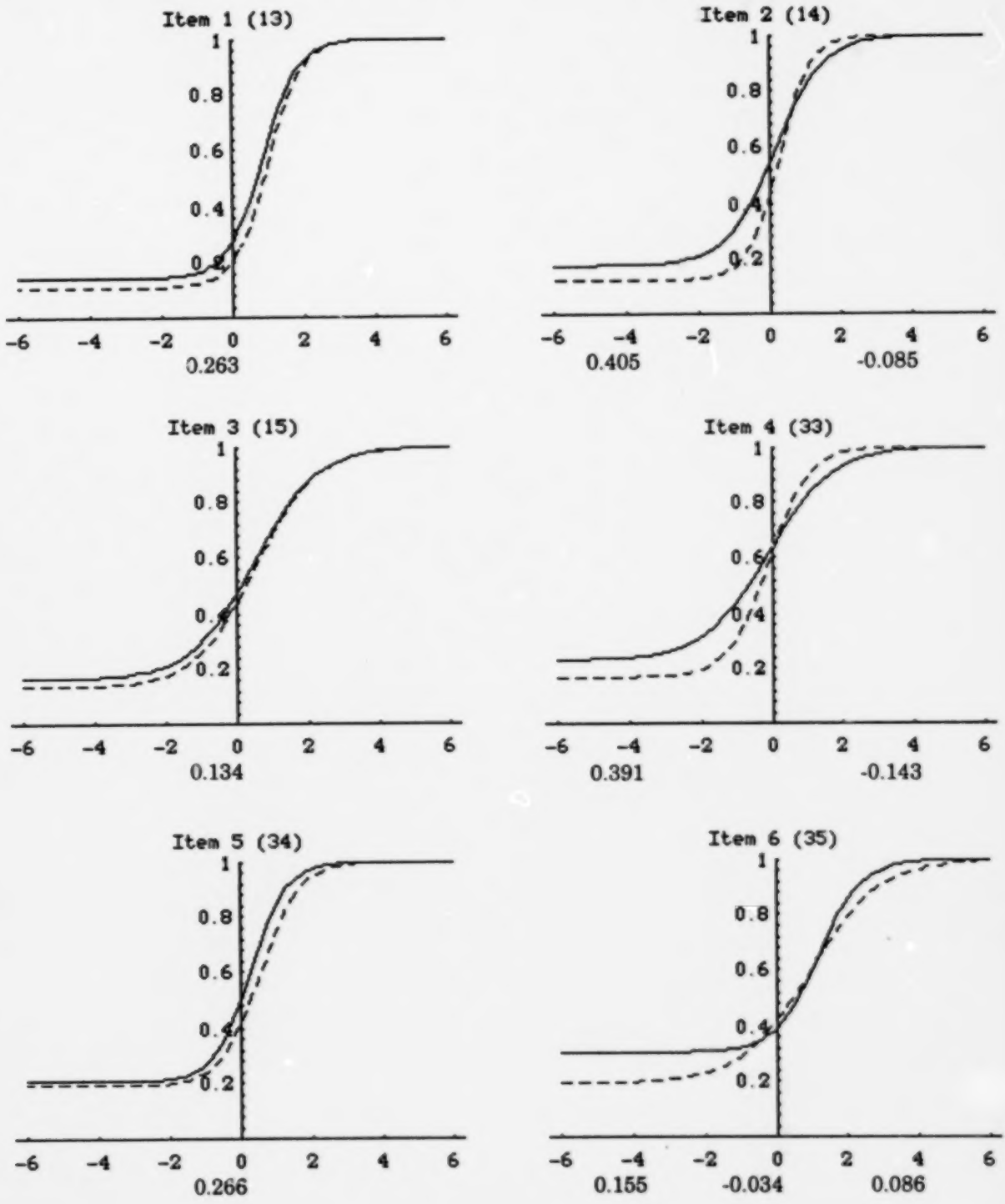


Figure 4. IRT-estimated item characteristic curves for 8th-grade algebra items (- - - - = nonalgebra group; —— = algebra group).

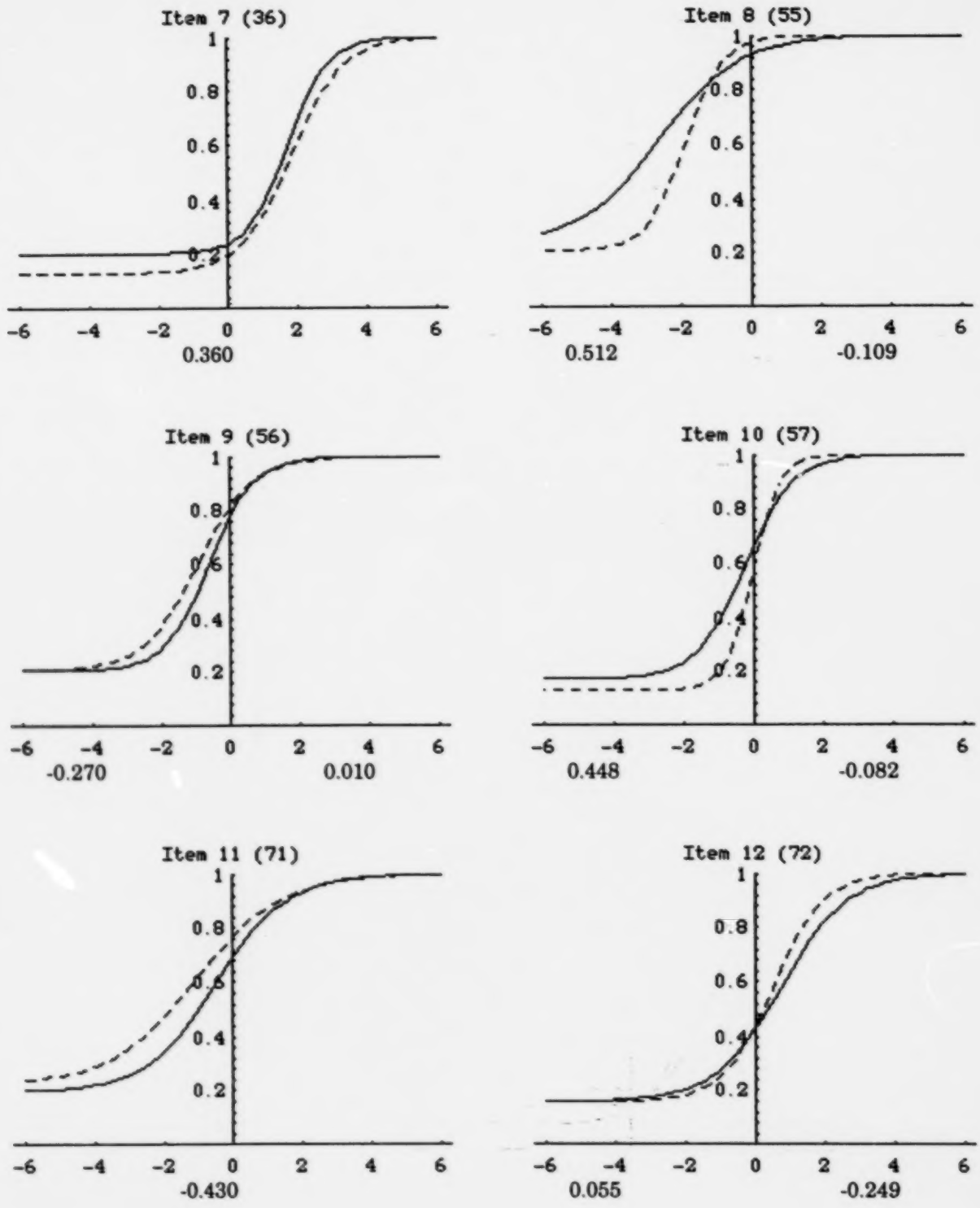
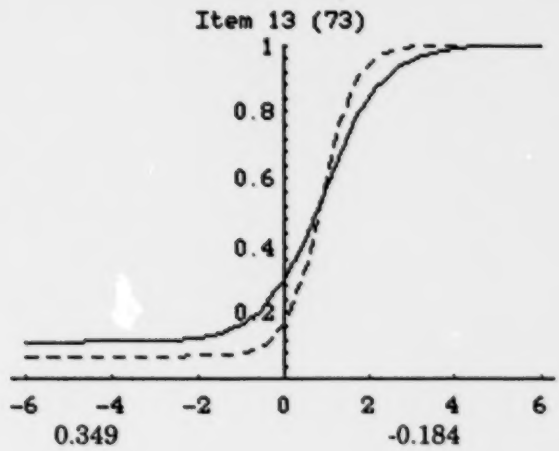
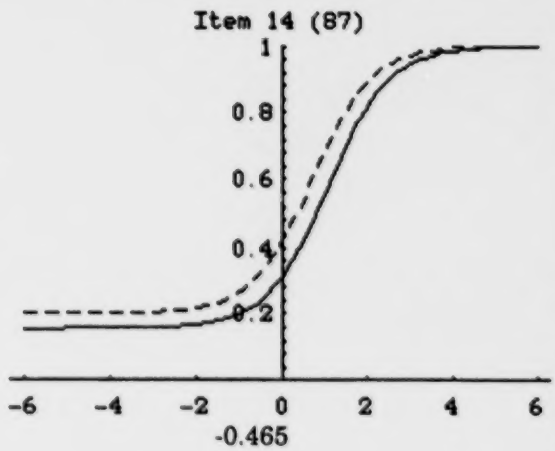
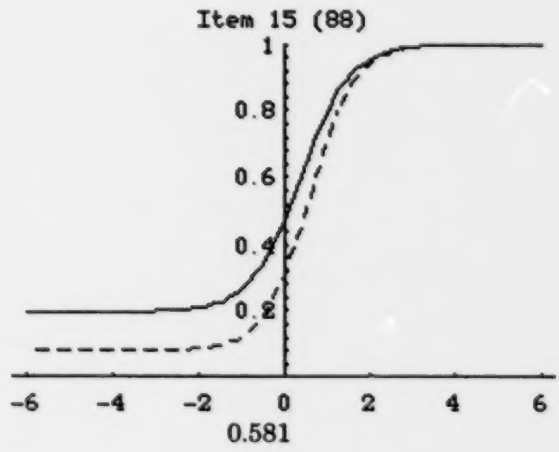
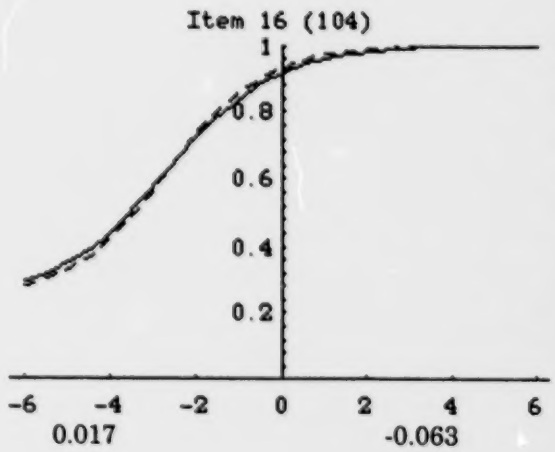


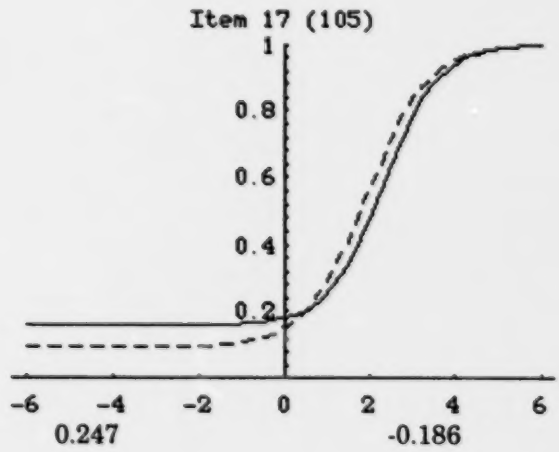
Figure 4. (continued)











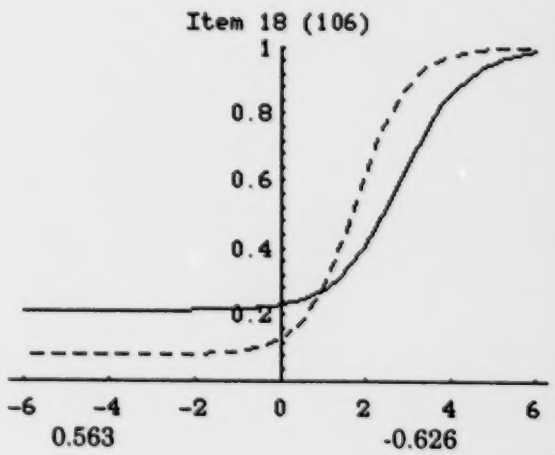
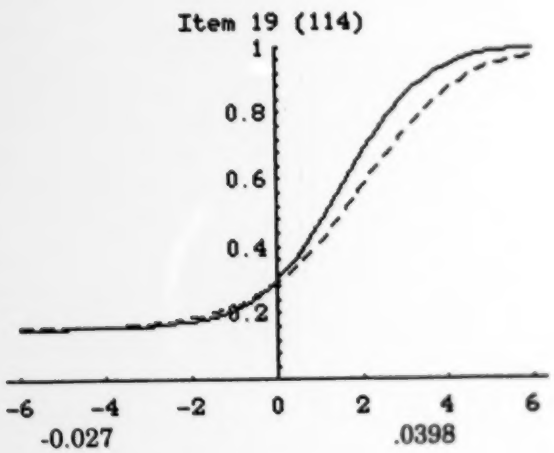
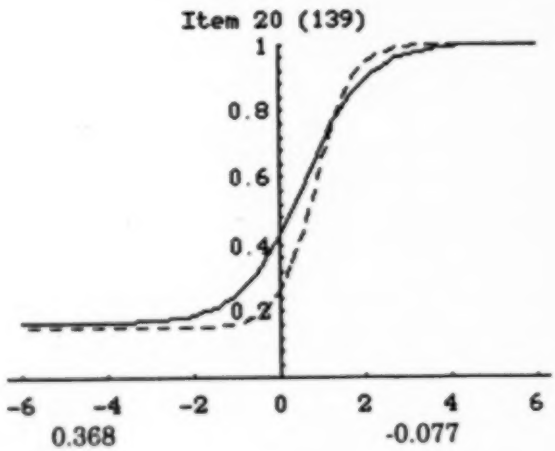
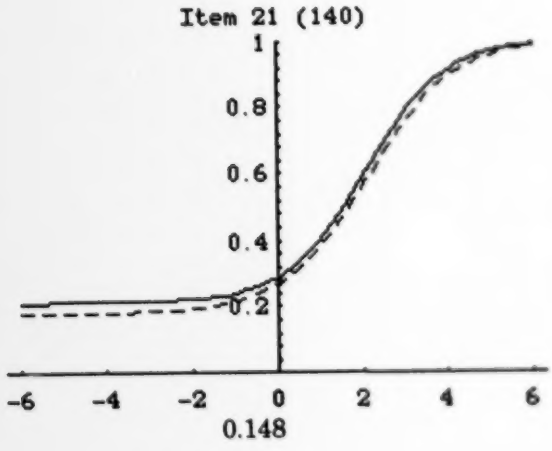
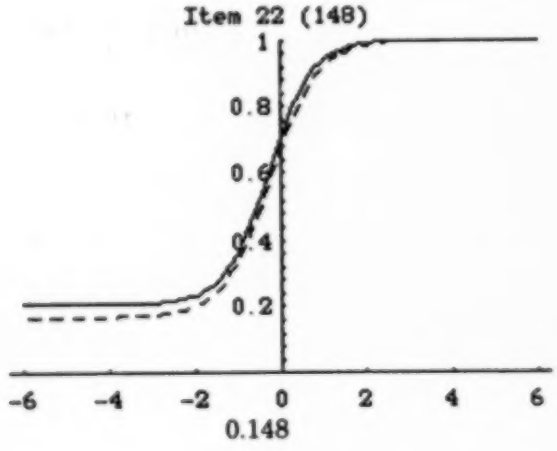


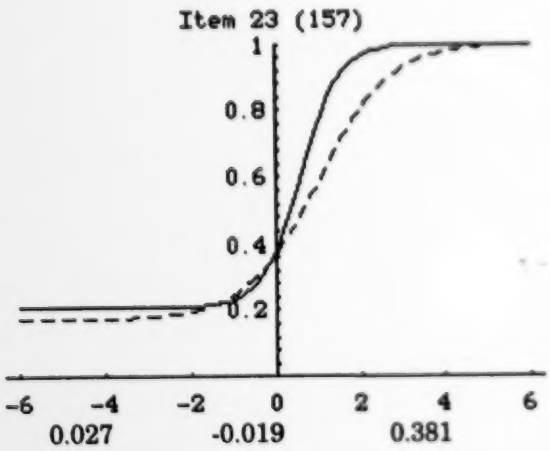
Figure 4. (continued)











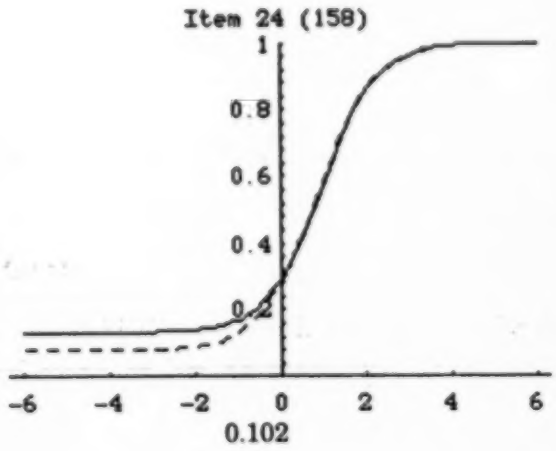


Figure 4. (continued)

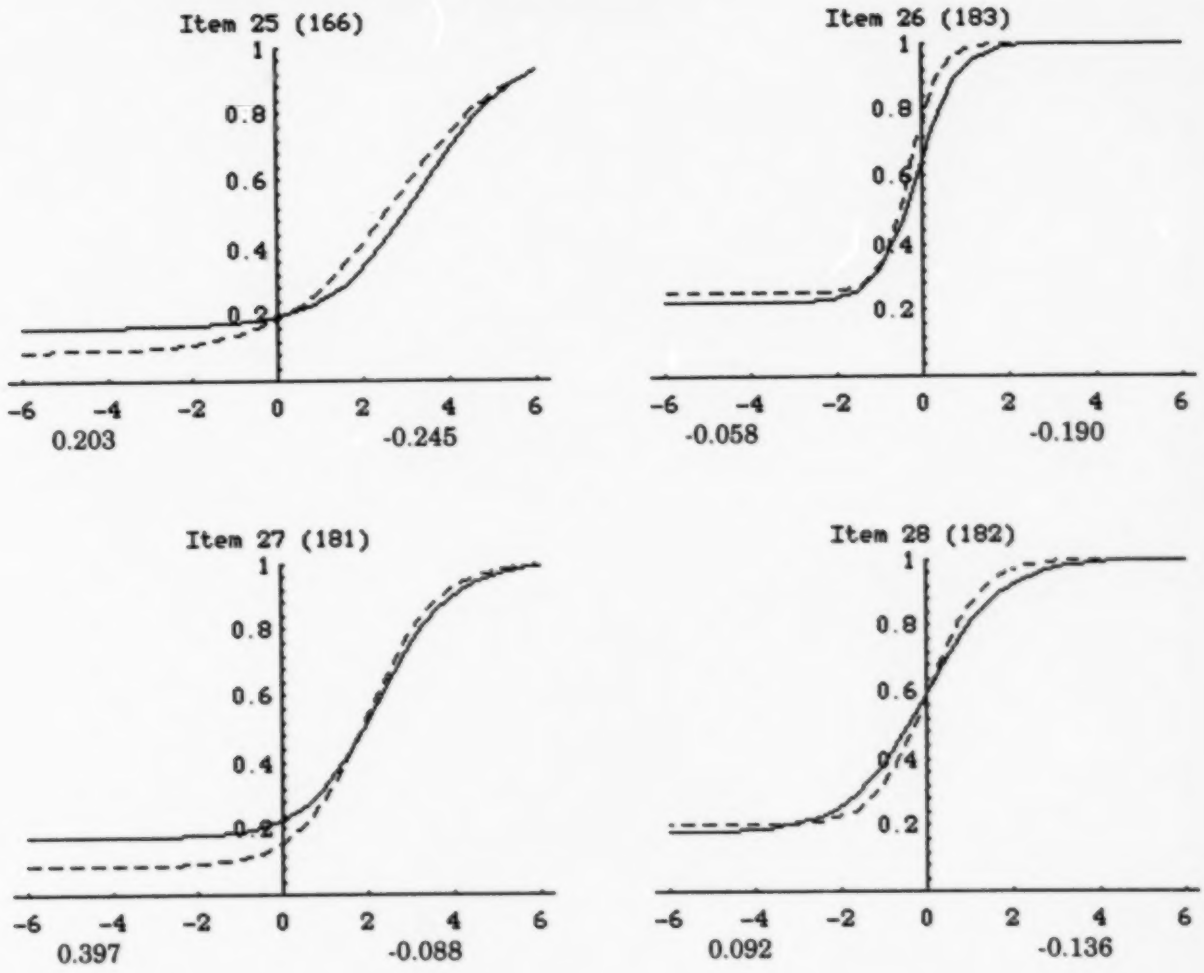


Figure 4. (continued)

Analyses of Test Items: Relating Extended Constructed-Response Items to OTL and Other Background Variables

Analysis Goals

The 1992 NAEP main math assessment for Grade 8 contained six extended constructed-response items which were rated on a graded scale corresponding to mathematical reasoning judged incorrect, minimal, partial, satisfactory, or extended. It is of interest to study these items in more detail with respect to OTL because they represent an item type that will presumably become more and more

common in NAEP and other assessments. Plans for the 1996 NAEP math test are that about 40% of the test will be open ended.

The goal is to relate the probability of doing well on each extended constructed-response item to the set of OTL variables discussed previously as well as to a set of key background variables that have proven to be important performance covariates in other analyses. It is of interest to see if the partial effect of OTL is significant, holding other factors constant. Note, however, that prior performance level is not controlled for. This analysis also serves as an indirect validity check for the OTL information in that good OTL measures should have an effect on the performance.

Methods

Logistic regression for an ordered categorical response variable is carried out for each extended constructed-response math item in the 1992 NAEP. The background variables used in these analyses are shown in Table 6.

Each analysis is carried out using a subset of the total sample based on data from 6 of the 26 booklets in which the item appears. Four of the six items had sufficient numbers of Grade 8 students in the partial or better categories and could be analyzed: item C—"reason to maximize difference" (numbers and operations); item I—"find probability and explain" (data analysis, statistics and probability); item L—"extend pattern to find term" (algebra); item M—"partition figures to find area" (measurement).

Results

Table 7 shows the results of the four logistic regressions. In the top panel the student-based algebra class type information is used. In the middle panel, teacher-based class type (remedial and typical, enriched, algebra) as well as the factor score for the "NCTM" factor of Table 1 are used. In the bottom panel both the student-based and teacher-based information are used.

Considering the top panel, student-reported algebra class membership has a significant influence on all items, as expected. Note again, however, that this effect is confounded with ability given the selection of algebra students based on prior performance. The teacher-based algebra class membership variable is also significant for all items, whereas the teacher-based enriched class type is significant only for item I (data analysis, probability and statistics) and item M

Table 6

Description of Background Variables for Analyses of Imputed Proficiencies (NAEP '92)

Sample size		4715 % in Grade 8	6293 % in Grade 12
OTL Student-Reported Class Type			
1. Algebra			
	*1 No Algebra/Other	50	
	2 Prealgebra	27	
	3 Algebra	23	
2. Alg-Calc			
	*1 Prealgebra/1st-Year Algebra/Not Studied		38
	2 2nd/3rd-Year Algebra		57
	3 Calculus		5
3. Geom-Trig			
	*1 Not Studied		21
	2 Geometry		58
	3 Trigonometry		21
OTL Teacher-Reported			
4. NCTM		—	
5. Class Type			
	*1 Remedial and Typical	37	
	2 Enriched	30	
6. School Program			
	*1 General		25
	2 Academic/College Prep		52
	3 Vocational/Technical		3
	3 Algebra	33	
7. Gender			
	*1 Male	49	46
	2 Female	51	54
8. Ethnicity			
	*1 White	71	72
	2 Black	16	16
	3 Hispanic	10	8
	4 Asian	3	4
9. Parents' Education (Student-reported)			
	1 Didn't Finish High School	8	7
	2 Grad From High School	25	21
	3 Some Ed After High School	20	27
	4 Grad From College	47	45

Table 6 (continued)

Sample size	4715 % in Grade 8	6293 % in Grade 12
10. Type of Community		
1 Extreme Rural	9	11
2 Disadvantaged Urban	8	12
3 Advantaged Urban	10	13
11. School Type		
*1 Public School	79	78
2 Private School	8	7
3 Catholic School	13	15
4 Other/Omitted		20
*4 Other (Non-Extreme)	73	64
12. TV Watching		
1 0-2 hours	41	
2 2-5 hours	47	
3 6 or more hours	12	
13. Language Other Than English (LOTE) at Home		
1 Never	66	
2 Sometimes	25	
3 Always	9	

Note. Categories in the background variables are all dummy coded except for Parents' Education, TV Watching, LOTE at Home and NCTM. For dummy-coded variables, effects are interpreted as the category in question compared to base category (marked *) of the variable.

(measurement). The latter finding is interesting given that enriched classes are characterized by an emphasis on measurement, geometry, and data analysis and statistics in line with the factor solution presented in Table 1. As shown in the bottom panel, these two effects are still significant when combining student-based and teacher-based information. This supports the validity of the teacher-based enriched class type derived from the factor analysis. The teacher-based algebra class type variable, however, appears to only add significant information beyond the student-based information for item M.

Further methodological approaches are possible on the item level when there is more detailed OTL information. As an example, the SIMS data provided item-specific OTL information which was utilized in analyses presented in Muthén (1994) and Muthén, Kao, and Burstein (1991).

Table 7 Items C, I, L, and M

		Item C: Numbers & Operations (reason to maximize difference) (N = 1077)				
		Parameter estimate	Standard error	Pr > Chi-square	Standardized estimate	Odds ratio
	Intercp1	-5.69	0.49	0.00	—	0.00
	Intercp2	-3.52	0.44	0.00	—	0.03
OTL Class Type (student rpt)	Prealgebra	0.49	0.24	0.04	0.12	1.64
	Algebra	1.53	0.23	0.00	0.35	4.62
Female		0.74	0.19	0.00	0.21	2.11
Ethnicity	Black	-0.92	0.37	0.01	-0.18	0.40
	Hispanic	-0.46	0.36	0.19	-0.09	0.63
	Asian	0.27	0.40	0.51	0.03	1.30
Parents' Ed.		0.28	0.11	0.02	0.15	1.32
Type of Community	Rural	0.16	0.34	0.64	0.03	1.17
	Disadv-Urban	-0.38	0.48	0.43	-0.06	0.68
	Adv-Urban	0.13	0.29	0.65	0.02	1.14
School Type	Private	-0.13	0.34	0.70	-0.02	0.88
	Catholic	-0.50	0.31	0.11	-0.09	0.61
	Intercp1	-5.60	0.48	0.00	—	0.00
	Intercp2	-3.47	0.44	0.00	—	0.03
OTL (teacher rpt)	NCTM	0.19	0.97	0.05	0.10	1.20
OTL Class Type (teacher rpt)	Enriched	-0.01	0.26	0.97	0.00	0.99
	Algebra	0.76	0.22	0.00	0.20	2.13
Female		0.67	0.19	0.00	0.18	1.95
Ethnicity	Black	-1.02	0.37	0.01	-0.20	0.36
	Hispanic	-0.60	0.35	0.09	-0.11	0.55
	Asian	0.32	0.40	0.42	0.03	1.38
Parents' Ed.		0.38	0.11	0.00	0.21	1.47
Type of Community	Rural	0.01	0.34	0.97	0.00	1.01
	Disadv-Urban	-0.22	0.47	0.64	-0.03	0.80
	Adv-Urban	0.21	0.29	0.45	0.04	1.24
School Type	Private	-0.10	0.34	0.76	-0.02	0.90
	Catholic	-0.63	0.30	0.04	-0.12	0.53
(Combined)						
	Intercp1	-5.68	0.50	0.00	—	0.00
	Intercp2	-3.51	0.45	0.00	—	0.03
OTL Class Type (student rpt)	Prealgebra	0.44	0.25	0.08	0.11	1.55
	Algebra	1.38	0.28	0.00	0.32	3.97
OTL (teacher rpt)	NCTM	0.13	0.10	0.16	0.07	1.14
OTL Class Type (teacher rpt)	Enriched	-0.03	0.26	0.92	-0.01	0.98
	Algebra	0.15	0.26	0.57	0.04	1.16
Female		0.74	0.19	0.00	0.21	2.10
Ethnicity	Black	-0.94	0.37	0.01	-0.18	0.39
	Hispanic	-0.48	0.36	0.18	-0.09	0.62
	Asian	0.20	0.41	0.63	0.02	1.22
Parents' Ed.		0.28	0.12	0.02	0.15	1.32
Type of Community	Rural	0.15	0.34	0.67	0.02	1.16
	Disadv-Urban	-0.35	0.48	0.46	-0.05	0.70
	Adv-Urban	0.11	0.29	0.72	0.02	1.11
School Type	Private	-0.10	0.34	0.78	-0.01	0.91
	Catholic	-0.50	0.31	0.11	-0.09	0.61

Table 7 (continued)

	Item I: Statistics and Probability (find probability and explain) (N = 1082)					
		Parameter estimate	Standard error	Pr > Chi-square	Standardized estimate	Odds ratio
	Intercp1	-3.75	0.41	0.00	—	0.02
	Intercp2	-2.78	0.39	0.00	—	0.06
OTL Class Type (student rpt)	Prealgebra	-0.11	0.25	0.66	-0.03	0.90
	Algebra	1.18	0.21	0.00	0.28	3.25
Female		-0.05	0.18	0.79	-0.01	0.95
Ethnicity	Black	-2.61	0.73	0.00	-0.50	0.07
	Hispanic	-0.61	0.34	0.07	-0.11	0.54
	Asian	0.38	0.40	0.35	0.03	1.46
Parents' Ed.		0.28	0.11	0.01	0.16	1.32
Type of Community	Rural	-0.16	0.39	0.68	-0.02	0.85
	Disadv-Urban	-0.44	0.50	0.39	-0.06	0.65
	Adv-Urban	0.87	0.24	0.00	0.15	2.38
School Type	Private	-0.30	0.33	0.36	-0.04	0.74
	Catholic	0.00	0.25	1.00	0.00	1.00
	Intercp1	-3.98	0.41	0.00	—	0.02
	Intercp2	-3.04	0.40	0.00	—	0.05
OTL (teacher rpt)	NCTM	0.21	0.09	0.02	0.12	1.23
OTL Class Type (teacher rpt)	Enriched	0.53	0.25	0.04	0.13	1.69
	Algebra	0.86	0.23	0.00	0.23	2.37
Female		-0.02	0.17	0.92	-0.01	0.98
Ethnicity	Black	-2.67	0.73	0.00	-0.51	0.07
	Hispanic	-0.68	0.33	0.04	-0.13	0.51
	Asian	0.54	0.40	0.17	0.05	1.71
Parents' Ed.		0.31	0.11	0.00	0.18	1.37
Type of Community	Rural	-0.22	0.39	0.57	-0.03	0.80
	Disadv-Urban	-0.22	0.50	0.66	-0.03	0.80
	Adv-Urban	0.90	0.24	0.00	0.15	2.45
School Type	Private	-0.40	0.33	0.22	-0.06	0.67
	Catholic	-0.14	0.25	0.57	-0.03	0.87
(Combined)						
	Intercp1	-3.92	0.43	0.00	—	0.02
	Intercp2	-2.95	0.41	0.00	—	0.05
OTL Class Type (student rpt)	Prealgebra	-0.13	0.25	0.60	-0.03	0.88
	Algebra	1.07	0.25	0.00	0.25	2.90
OTL (teacher rpt)	NCTM	0.13	0.09	0.16	0.07	1.14
OTL Class Type (teacher rpt)	Enriched	0.49	0.25	0.05	0.13	1.64
	Algebra	0.34	0.26	0.20	0.09	1.40
Female		-0.05	0.18	0.76	-0.02	0.95
Ethnicity	Black	-2.63	0.73	0.00	-0.50	0.07
	Hispanic	-0.66	0.34	0.05	-0.12	0.52
	Asian	0.37	0.40	0.36	0.03	1.45
Parents' Ed.		0.26	0.11	0.02	0.15	1.29
Type of Community	Rural	-0.09	0.39	0.82	-0.01	0.91
	Disadv-Urban	-0.37	0.50	0.47	-0.05	0.69
	Adv-Urban	0.90	0.25	0.00	0.15	2.46
School Type	Private	-0.29	0.33	0.39	-0.04	0.75
	Catholic	-0.04	0.25	0.87	-0.01	0.96

Table 7 (continued)

		Item L: Algebra (extend pattern to find term) (N = 1090)				
		Parameter estimate	Standard error	Pr > Chi-square	Standardized estimate	Odds ratio
	Intercp1	-4.75	0.62	0.00	—	0.01
	Intercp2	-4.57	0.62	0.00	—	0.01
OTL Class Type (student rpt)	Prealgebra	0.57	0.34	0.10	0.14	1.77
	Algebra	1.71	0.30	0.00	0.39	5.55
Female		0.20	0.24	0.41	0.06	1.22
Ethnicity	Black	-0.86	0.54	0.11	-0.16	0.43
	Hispanic	-0.75	0.62	0.23	-0.14	0.47
	Asian	-0.89	0.68	0.19	-0.08	0.41
Parents' Ed.		0.33	0.16	0.03	0.18	1.40
Type of Community	Rural	0.55	0.45	0.22	0.08	1.73
	Disadv-Urban	-0.24	0.76	0.75	-0.03	0.79
	Adv-Urban	0.52	0.32	0.11	0.09	1.68
School Type	Private	0.86	0.36	0.02	0.13	2.36
	Catholic	0.38	0.33	0.26	0.07	1.46
	Intercp1	-4.69	0.60	0.00	—	0.01
	Intercp2	-4.52	0.60	0.00	—	0.01
OTL (teacher rpt)	NCTM	-0.07	0.12	0.58	-0.04	0.94
OTL Class Type (teacher rpt)	Enriched	0.05	0.36	0.90	0.01	1.05
	Algebra	0.97	0.30	0.00	0.25	2.63
Female		0.13	0.24	0.57	0.04	1.14
Ethnicity	Black	-0.90	0.54	0.09	-0.17	0.41
	Hispanic	-0.86	0.62	0.16	-0.16	0.42
	Asian	-0.50	0.66	0.44	-0.04	0.61
Parents' Ed.		0.44	0.15	0.00	0.24	1.55
Type of Community	Rural	0.46	0.44	0.29	0.07	1.59
	Disadv-Urban	-0.23	0.76	0.77	-0.03	0.80
	Adv-Urban	0.57	0.31	0.07	0.10	1.76
School Type	Private	0.70	0.34	0.04	0.11	2.01
	Catholic	0.24	0.33	0.45	0.04	1.28
(Combined)						
	Intercp1	-4.87	0.64	0.00	—	0.01
	Intercp2	-4.69	0.64	0.00	—	0.01
OTL Class Type (student rpt)	Prealgebra	0.52	0.35	0.14	0.13	1.68
	Algebra	1.62	0.35	0.00	0.37	5.03
OTL (teacher rpt)	NCTM	-0.15	0.12	0.23	-0.08	0.86
OTL Class Type (teacher rpt)	Enriched	0.05	0.37	0.90	0.01	1.05
	Algebra	0.31	0.34	0.35	0.08	1.37
Female		0.18	0.24	0.45	0.05	1.20
Ethnicity	Black	-0.79	0.54	0.14	-0.15	0.45
	Hispanic	-0.71	0.63	0.26	-0.13	0.49
	Asian	-0.89	0.67	0.19	-0.08	0.41
Parents' Ed.		0.34	0.16	0.03	0.19	1.41
Type of Community	Rural	0.64	0.45	0.16	0.10	1.89
	Disadv-Urban	-0.27	0.76	0.72	-0.04	0.76
	Adv-Urban	0.48	0.33	0.15	0.08	1.61
School Type	Private	0.84	0.36	0.02	0.13	2.32
	Catholic	0.33	0.33	0.32	0.06	1.39

Table 7 (continued)

		Item M: Measurement (partition figures to find area) (N = 1142)				
		Parameter estimate	Standard error	Pr > Chi-square	Standardized estimate	Odds ratio
	Intercp1	-5.18	0.63	0.00	—	0.01
	Intercp2	-4.68	0.62	0.00	—	0.01
OTL Class Type (student rpt)	Prealgebra	0.42	0.35	0.24	0.10	1.52
	Algebra	1.73	0.30	0.00	0.40	5.65
Female		0.17	0.24	0.48	0.05	1.18
Ethnicity	Black	-1.80	0.74	0.02	-0.34	0.17
	Hispanic	-0.89	0.55	0.11	-0.18	0.41
	Asian	0.87	0.38	0.02	0.10	2.37
Parents' Ed.		0.40	0.16	0.01	0.23	1.49
Type of Community	Rural	-0.07	0.51	0.90	-0.01	0.94
	Disadv-Urban	0.43	0.51	0.40	0.07	1.54
	Adv-Urban	0.60	0.32	0.06	0.10	1.81
School Type	Private	0.32	0.37	0.38	0.05	1.38
	Catholic	0.02	0.35	0.95	0.01	1.02
	Intercp1	-5.64	0.65	0.00	—	0.00
	Intercp2	-5.16	0.64	0.00	—	0.01
OTL (teacher rpt)	NCTM	0.12	0.12	0.32	0.07	1.13
OTL Class Type (teacher rpt)	Enriched	1.12	0.39	0.00	0.28	3.05
	Algebra	1.52	0.36	0.00	0.39	4.56
Female		0.24	0.23	0.30	0.07	1.28
Ethnicity	Black	-1.99	0.74	0.01	-0.37	0.14
	Hispanic	-1.11	0.55	0.04	-0.22	0.33
	Asian	0.93	0.37	0.01	0.10	2.54
Parents' Ed.		0.44	0.16	0.01	0.25	1.55
Type of Community	Rural	0.18	0.51	0.72	0.03	1.20
	Disadv-Urban	0.67	0.51	0.19	0.10	1.95
	Adv-Urban	0.71	0.31	0.02	0.12	2.02
School Type	Private	0.48	0.36	0.18	0.07	1.62
	Catholic	-0.13	0.34	0.72	-0.02	0.88
(Combined)						
	Intercp1	-5.72	0.68	0.00	—	0.00
	Intercp2	-5.21	0.67	0.00	—	0.01
OTL Class Type (student rpt)	Prealgebra	0.44	0.36	0.22	0.11	1.56
	Algebra	1.57	0.34	0.00	0.36	4.81
OTL (teacher rpt)	NCTM	0.04	0.13	0.77	0.02	1.04
OTL Class Type (teacher rpt)	Enriched	1.06	0.40	0.01	0.27	2.89
	Algebra	0.86	0.38	0.03	0.22	2.36
Female		0.20	0.24	0.41	0.06	1.22
Ethnicity	Black	-1.80	0.74	0.02	-0.34	0.17
	Hispanic	-0.96	0.55	0.08	-0.19	0.38
	Asian	0.88	0.38	0.02	0.10	2.40
Parents' Ed.		0.35	0.16	0.03	0.20	1.42
Type of Community	Rural	0.18	0.52	0.73	0.03	1.20
	Disadv-Urban	0.46	0.52	0.37	0.07	1.59
	Adv-Urban	0.65	0.33	0.05	0.11	1.91
School Type	Private	0.37	0.37	0.33	0.05	1.44
	Catholic	-0.02	0.35	0.96	0.00	0.98

Analyses of Multivariate Scores in Relation to OTL

Analysis Goals

We will now return to the issue of controlling for prior performance when attempting to assess OTL effects on performance. The analysis is now on the score level. Our idea is to use a general math factor that influences performance in all content areas. As mentioned earlier, the introduction of such a general factor notion has two advantages: OTL effects can be clearly described as effects that go beyond what is expected by the general factor, and it is more likely that interesting OTL effects can be found with respect to performance that is over and above that expected by a general factor. Three types of analyses are of interest: analysis of NAEP proficiency scores, analysis of NAEP process scores, and longitudinal analysis of NELS data.

Multivariate analysis of NAEP proficiency scores. The first analysis concerns the NAEP-provided proficiency scores, which are imputed in five versions for each of five math content areas. A latent variable model is of interest for these proficiency scores, where a general and several specific factors can be identified and regressed on background information including OTL. In the main NAEP assessment, proficiency scores are produced for each of the five content areas: numbers and operations, measurement, geometry, data analysis and statistics, and algebra. We will formulate a multivariate response model for these five areas and relate it to the same set of OTL and background variables as studied earlier. The response model is formulated with the aim of separating out a general factor influencing performance on all content areas from specific factors influencing only one content area.

The partial effects of OTL, given other background variables, on the general factor often correspond to ability differences due to selection effects as with 8th-grade tracking for algebra classes. OTL effects on the content-specific factors, however, correspond perhaps more closely to the question initially posed, "What students know and can do as a result of their educational experiences." The main interest of an OTL analysis is in studying the partial OTL effect on the specific factors, where effects of the general factor as well as other background variables are held constant.

If one can assume that in comparison to skills in each content area, the general skills are relatively stable from, say, Grade 7 to Grade 8, the general

BEST COPY AVAILABLE

factor provides a reasonable proxy for prior performance. This is a hypothesis that needs to be tested, however, perhaps using longitudinal data such as NELS. If this is found to be a reasonable approximation, the general factor would become an important variable to control for when studying OTL effects on performance in particular content areas.

Multivariate analysis of NAEP process scores. A similar latent variable analysis is applied to scores related to processes or categories of mathematical abilities. The NAEP 1992 Technical Report (Johnson & Carlson, 1994), refers to three mathematical ability categories within which the math items can be organized: conceptual understanding, procedural knowledge, and problem solving (see page 52). Given the NCTM *Standards*, it is of particular interest to study to which extent a problem-solving factor can be identified. A discussion of the NAEP 1990 Grade 8 math items in the context of NCTM standards is given in Silver and Kenney (1993). NAEP 1992 defines problem-solving items as follows (Johnson & Carlson, 1994):

> In problem solving, students are required to use their reasoning and analytic abilities when they formulate problems; determine the sufficiency and consistency of the data; use strategies, data models, and relevant mathematics; generate, extend, and modify procedures; use reasoning (i.e., spatial, inductive, deductive, statistical, and proportional); and judge the reasonableness and correctness of solutions. (p. 52)

Longitudinal analysis of NELS first follow-up data. Longitudinal achievement data have the potential of more clearly disentangling effects of OTL from effects of prior performance level. Using the first follow-up NELS data from Grades 8 and 10, it is of interest to model 10th-grade performance as a function of both 8th-grade and 10th-grade OTL while conditioning on 8th-grade math ability. In this article reading performance is also available as a covariate.

Methods

For all three analyses, structural equation modeling with latent variables and maximum-likelihood estimation will be used.

Multivariate analysis of NAEP's proficiency scores. The latent variable model is shown in diagram form in Figure 5. In it, the general factor is similar to NAEP's overall math performance score and to a large extent represents skills related to arithmetic, particularly as represented by the numbers and operations items. Specific factors represent student variation in content-area scores, which

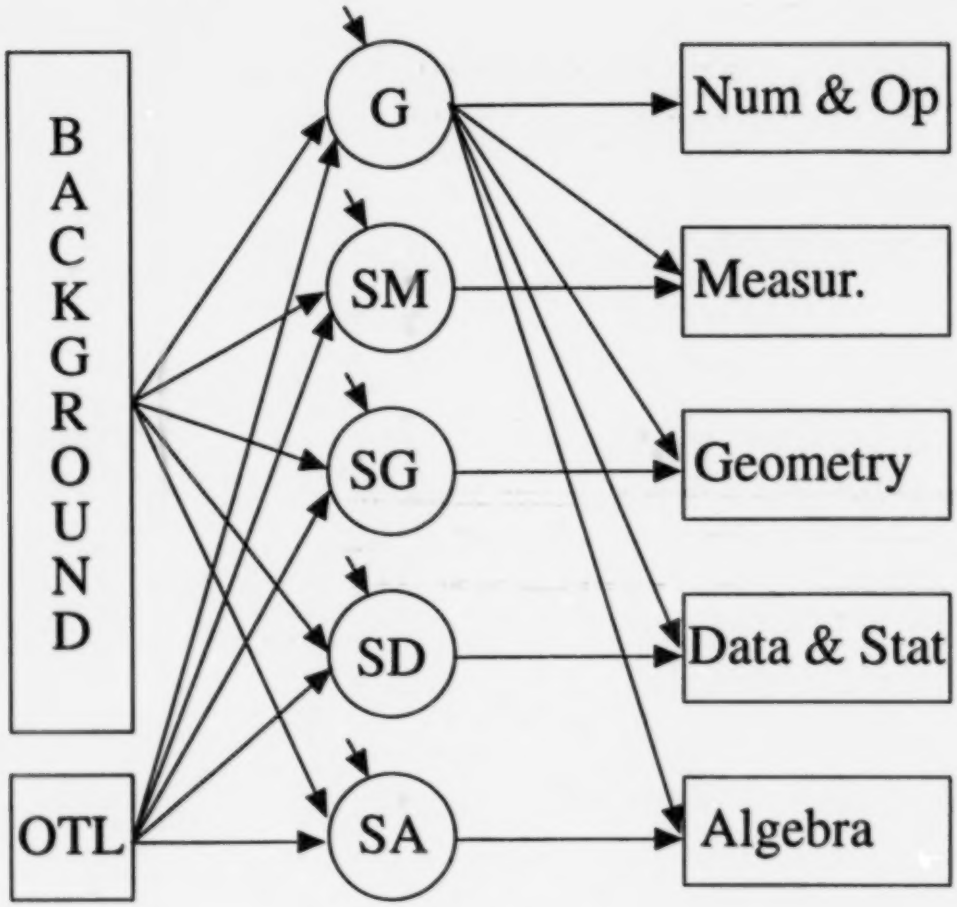


Figure 5. **Multivariate NAEP model using multiple imputations from five math content areas.**

differ from that of general arithmetic tasks and are more directly related to topics, definitions, and skills specific to the content area. A specific factor is not included for numbers and operations so that the general factor is more clearly defined in terms of such arithmetic skills. This type of modeling was successfully applied to NAEP math data in Muthén, Khoo, and Goff (1994) to demonstrate multidimensionality in math performance. In that application, however, specially derived testlets were used corresponding to item content and format, whereas in the present case the usual content-specific proficiencies produced by NAEP are used. Using the model of Figure 5, the performance in a certain content area is

BEST COPY AVAILABLE

decomposed into effects from a general factor and a content-specific factor, and each of these two factors is related to OTL and other background.

The analysis of the Figure 5 model is carried out using the five multiple imputation values that are provided in NAEP for each of the five content areas. In comparison to the true ("theta") proficiency score for each of the five content areas, these imputed proficiency values have been created such that they have the same mean, variance, and covariance with background variables (Mislevy, 1991, 1993), as well as having the same covariances between content areas (Mislevy, personal communication, August 1994). A latent variable regression model for the true proficiency scores can therefore be analyzed using the imputed values as observed variables. This implies that there are no measurement error components for the observed variables in the model of Figure 5. The residuals for the specific factors are defined to be uncorrelated with the residual for the general factor. The model has six degrees of freedom if the specific factor residuals are taken to be uncorrelated among themselves, although this is not a necessary assumption for identifying the model. The input for structural equation modeling software is given in the appendix. Input for the LISCOMP program is used here, but LISREL input is very similar.

As is customary with multiple imputations, the model is analyzed for each imputed value, and the estimates are averaged over the five analyses. Standard errors of estimates are calculated by the usual imputation formula using the combination of average standard error over imputations and between-imputation estimate variability (see, e.g., Mislevy, Johnson, & Muraki, 1992). The "grade only" main 1992 NAEP samples are here analyzed for Grade 8 and Grade 12. In each case, the student-based and the teacher-based OTL variables considered earlier are used together.

Multivariate analysis of process scores in NAEP. These scores are not provided by NAEP but are generated as parcels of items classified into the NAEP ability categories conceptual understanding, procedural knowledge, and problem solving. For this analysis, three types of testlets were formed corresponding to the three ability categories, ignoring math content. The item classification was obtained from ETS (Pashley, personal communication, 6 October 1994). The testlets were formed for items in each block for all of the 26 booklets. On average, six testlets per block were formed with most testlets consisting of two items.

There were 29 testlets for conceptual understanding, 29 for problem solving, and 17 for procedural knowledge.

A multifactorial model in line with Muthén et al. (1994) was used also here. The analyses reported here are limited to considering a problem-solving factor in addition to a general factor. The general factor was defined by having procedural knowledge and conceptual understanding variables load only on this. The problem-solving variables, however, were allowed to load not only on the general factor but also on a specific problem-solving factor. In line with Muthén et al. (1994), a simultaneous structural modeling analysis was carried out for the 26 groups of students defined by the 26 NAEP booklets. The same set of background variables as shown in Table 6 were used. Here, results will only be reported for Grade 8.

Longitudinal analysis of NELS first follow-up data. The model is shown in Figure 6. The analysis was carried out on testlets created from the items. An algebra-specific factor was defined in addition to a general factor at both Grade 8 and 10. Invariance of factor loadings across the grades was not imposed, allowing for changes in measurement characteristics over time. At Grade 10, three mathematics test forms were used depending on the math performance in Grade 8. Here, 2,413 students from the middle group and 1,418 students from the high group were included in the analysis. A simultaneous analysis of these two groups was carried out. To reflect the fact that the analysis pertains to a single population, parameters representing the same quantities were held equal across the two groups.

Results

Multivariate analysis of NAEP's proficiency scores. The latent variable model estimates are shown in Tables 8–11. The analyses showed that the results were almost exactly the same whether the specific factor residuals were specified as uncorrelated or not, although nonzero correlations were indicated (for Grade 8, for example, the residual correlations range from .23 for data analysis and statistics and measurement to .46 for geometry and measurement as well as for algebra and geometry, while for Grade 12 the residual correlations range from .20 for data analysis and statistics and measurement to .64 for geometry and measurement). Because the model with correlated residuals is just identified, no chi-square measure of model fit is provided.

BEST COPY AVAILABLE

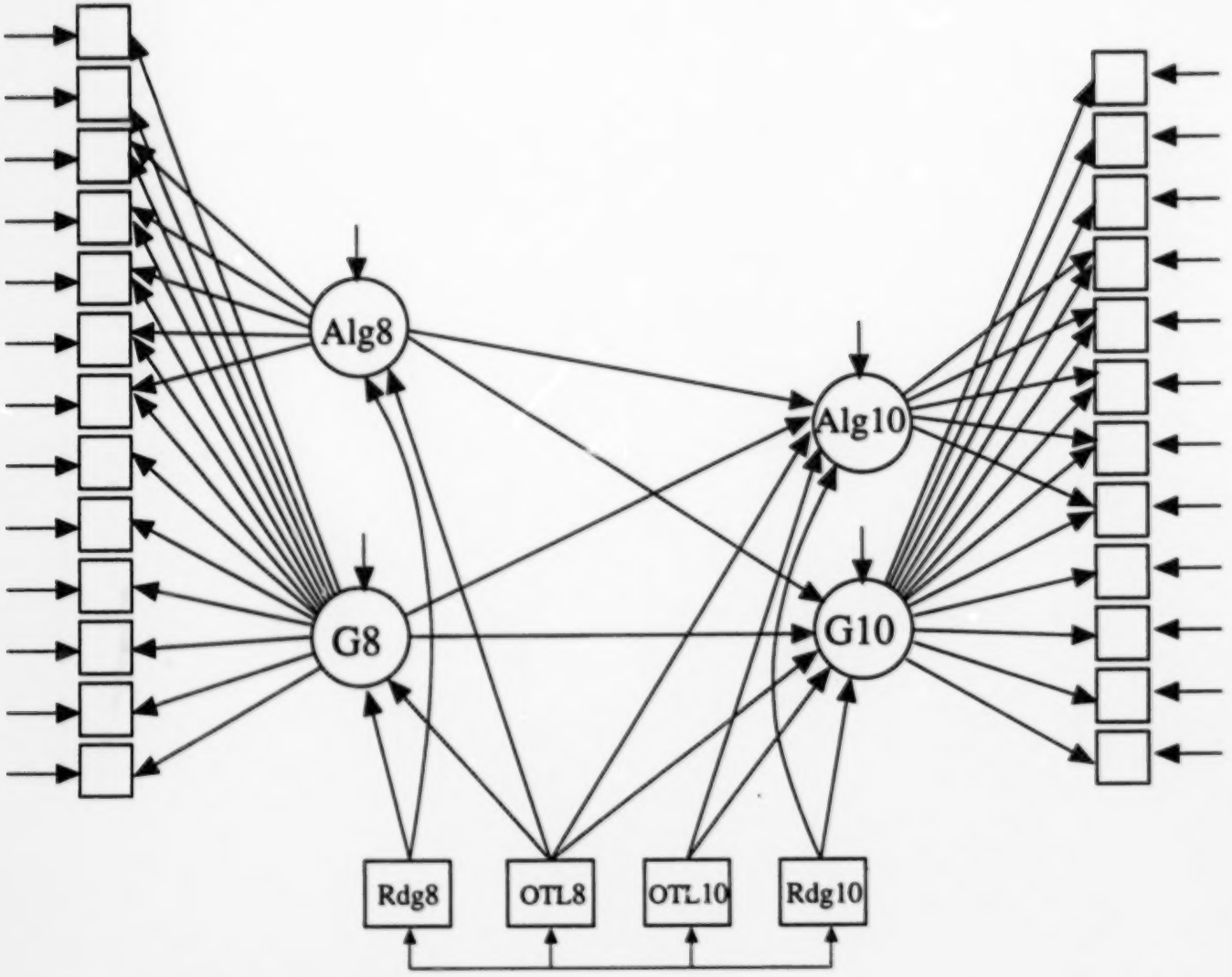


Figure 6. Longitudinal model for NELS.

37

38

Table 8. NAEP 1992 Grade 8 Factor Loadings (Standard Errors) From the Structural Model for Imputed Proficiencies Regressing One General and Four Specific Factors on 17 x-Variables (n=4715)

	Factors				
	General	Meas.	Geom.	Data & Stat	Algebra
Proficiencies					
Numbers & Operations	0.90 (0.02)				
Measurement	1.00 —	1.00 —			
Geometry	0.75 (0.02)		1.00 —		
Data & Statistics	0.99 (0.02)			1.00 —	
Algebra	0.86 (0.01)				1.00 —
Percentage of proficiency variance explained by specific factors		17	24	10	10

Note. Empty entries correspond to loadings fixed at zero.

Table 9. NAEP '92 Grade 12 Factor Loadings (Standard Errors) From the Structural Model for Imputed Proficiencies Regressing One General and Four Specific Factors on 17 x-Variables (n=6293)

	Factors				
	General	Meas.	Geom.	Data & Stat	Algebra
Proficiencies					
Numbers & Operations	1.01 (0.01)				
Measurement	1.00 —	1.00 —			
Geometry	0.97 (0.02)		1.00 —		
Data & Statistics	0.96 (0.02)			1.00 —	
Algebra	0.92 (0.01)				1.00 —
Percentage of proficiency variance explained by specific factors		14	20	12	16

Note. Empty entries correspond to loadings fixed at zero.

BEST COPY AVAILABLE

Table 10

NAEP 1992 Grade 8 Standardized Coefficients (*t*-values) From the Structural Model Regressing One General and Four Specific Factors on 17 x-Variables (*n*=4715)

	General	Meas.	Geom.	Data	Algebra
OTL (Student-Reported Class Type)					
Prealgebra	0.17 *5.15*	-0.09 *-1.37*	0.03 *0.34*	0.07 *0.44*	0.07 *0.91*
Algebra	0.67 *17.86*	0.12 *1.53*	0.08 *1.38*	0.09 *1.03*	0.57 *4.66*
(Teacher-Reported)					
NCTM	0.00 *-0.02*	0.06 *1.98*	0.03 *0.89*	0.02 *0.45*	0.15 *3.24*
Class Type					
Enriched	0.19 *6.05*	0.08 *1.08*	-0.02 *-0.39*	-0.10 *-1.76*	-0.05 *-0.47*
Algebra	0.36 *10.31*	0.02 *0.22*	0.05 *0.71*	-0.03 *-0.23*	0.12 *1.14*
Female	0.00 *-0.08*	-0.17 *-8.54*	-0.04 *-1.46*	-0.04 *-1.55*	0.06 *1.53*
Ethnicity					
Black	-0.73 *-20.43*	-0.49 *-6.47*	-0.39 *-4.30*	-0.38 *-6.27*	-0.29 *-2.11*
Hispanic	-0.34 *-6.55*	0.06 *1.02*	-0.01 *-0.10*	-0.23 *-1.75*	-0.24 *-2.55*
Asian	0.31 *3.58*	-0.03 *-0.17*	-0.08 *-0.62*	-0.39 *-2.72*	0.01 *0.04*
Parents' Ed.	0.15 *9.56*	0.07 *1.70*	0.03 *0.78*	0.10 *1.74*	0.06 *1.40*
Type of Community					
Rural	0.09 *1.89*	-0.05 *-0.71*	-0.22 *-3.31*	-0.04 *-0.22*	0.03 *0.22*
Disadv-Urban	-0.29 *-5.21*	-0.20 *-1.41*	-0.03 *-0.25*	-0.06 *-0.43*	0.17 *1.41*
Adv-Urban	0.18 *4.23*	0.15 *2.20*	0.18 *3.22*	0.07 *0.48*	0.32 *2.55*
School Type					
Catholic	0.08 *2.08*	-0.13 *-1.27*	-0.10 *-1.31*	-0.17 *-1.30*	-0.18 *-2.27*
Private	0.07 *1.20*	-0.21 *-2.08*	-0.12 *-1.31*	-0.07 *-0.59*	-0.18 *-0.92*
TV Watching	-0.09 *-6.66*	-0.09 *-2.79*	-0.03 *-1.93*	-0.07 *-2.86*	0.01 *0.43*
Home Language	-0.04 *-2.82*	-0.03 *-0.86*	0.00 *0.09*	-0.02 *-0.58*	0.00 *0.06*
R Square	0.39	0.11	0.05	0.08	0.17
Proficiency Variance [explained by specific factors]	—	0.17	0.24	0.10	0.10

Table 11

NAEP 1992 Grade 12 Standardized Coefficients (*t*-values) From the Structural Model Regressing One General and Four Specific Factors on 17 x-Variables (*n*=6293)

	General	Meas.	Geom.	Data	Algebra
OTL					
Alg-Calc					
Algebra	0.34 *13.33*	0.16 *2.84*	0.17 *3.55*	0.01 *0.25*	0.45 *12.34*
Calculus	0.88 *10.37*	0.42 *1.74*	0.34 *2.17*	0.05 *0.39*	0.71 *6.72*
Geom-Trig					
Geometry	0.45 *11.77*	0.20 *2.40*	0.60 *10.96*	0.18 *3.18*	0.62 *6.11*
Trigonometry	0.49 *9.29*	0.27 *2.19*	0.75 *8.52*	0.18 *2.04*	0.91 *8.66*
School-Program					
Academic	0.43 *15.39*	0.09 *1.58*	0.14 *3.61*	-0.02 *-0.46*	0.30 *9.35*
Vocational	-0.05 *-0.64*	0.22 *2.21*	-0.07 *-0.72*	-0.09 *-0.35*	-0.02 *-0.12*
Other	0.05 *1.43*	0.06 *0.94*	0.14 *3.01*	-0.16 *-1.93*	0.15 *2.54*
Female	-0.06 *-6.06*	-0.16 *-5.37*	-0.11 *-4.75*	-0.03 *-0.62*	0.00 *0.19*
Ethnicity					
Black	-0.65 *-19.54*	-0.49 *-4.91*	-0.06 *-1.06*	-0.35 *-3.98*	0.07 *1.33*
Hispanic	-0.28 *-4.91*	0.16 *1.19*	0.32 *2.77*	0.05 *0.65*	0.17 *2.32*
Asian	0.05 *0.61*	0.46 *2.24*	0.43 *5.17*	-0.40 *-2.69*	0.51 *3.91*
Parents' Ed.	0.11 *8.97*	0.08 *3.18*	0.07 *3.00*	0.07 *2.07*	0.04 *1.71*
Type of Community					
Rural	-0.05 *-1.32*	-0.10 *-1.63*	-0.11 *-1.92*	-0.04 *-0.42*	-0.17 *-2.27*
Disadv-Urban	-0.15 *-4.22*	-0.16 *-3.00*	-0.03 *-0.57*	-0.05 *-0.72*	-0.05 *-0.96*
Adv-Urban	0.15 *3.34*	0.18 *2.71*	0.17 *1.38*	0.20 *0.96*	0.16 *2.19*
School Type					
Catholic	-0.10 *-2.73*	0.03 *0.26*	-0.06 *-0.56*	-0.02 *-0.34*	0.10 *1.37*
Private	0.20 *3.75*	0.05 *0.45*	0.03 *0.25*	-0.07 *-0.46*	0.19 *1.96*
R Square	0.41	0.17	0.19	0.07	0.36
Proficiency Variance [explained by specific factors]	—	0.14	0.20	0.12	0.16

BEST COPY AVAILABLE

Tables 8 and 9 show the measurement part of the model for Grades 8 and 12 (cf. Figure 5). With the exception of the numbers and operations proficiency, all five observed variables load on two factors, the general and a specific factor. The factor variances are estimated but not shown here. To facilitate understanding, the information on these values is instead given as the percentage variance that each specific factor accounts for in the corresponding observed proficiency variable. It is, for example, seen that the geometry proficiency has the lowest loading on the general factor and has the largest percentage of its variance explained by the geometry-specific factor.

Consider next the Table 10 structural modeling results for Grade 8. The first column contains the effects of the background variables including OTL on the general factor. Strong effects of algebra OTL are seen on the general factor, to a large extent reflecting the selection of students into such classes based on prior performance. Student-based algebra class type has a large effect, while the effect of prealgebra is considerably smaller. This is in line with the item performance shown earlier in Figures 2 and 3, where the difference between the solid and the broken line, describing the average math item performance difference of students in (pre) algebra classes and nonalgebra classes, was large for the algebra group but not for the prealgebra group. We note that the teacher-based class type variable has effects on the general factor beyond that of the student-based OTL information.

The results for the specific factors are of particular interest. There is an especially large effect of student-based algebra class type on the algebra-specific factor. This is in line with Figure 2 where many algebra items were found above the solid line describing the average item performance of students in algebra classes. With the exception of the effects of the teacher-reported NCTM factor and algebra class type on the algebra-specific factor, however, the remaining OTL variables do not show any effects. In particular, the absence of effects from the teacher-based enriched class type on the content-specific factors measurement, geometry, and data analysis and statistics is noteworthy. Either this class type variable is not well enough defined by the teacher-based emphasis variables or the increased emphasis is not sufficiently relevant for these test items.

Finally, it is of interest to consider how this table describes ethnicity differences in performance. This may be related to the introductory discussion in connection with algebra performance shown in Figure 1. In line with Figure 1, the

results in the general column of Table 10 indicate that there are still large ethnicity differences for the general factor even when OTL is taken into account. In several instances, the content-specific columns also show significant ethnicity coefficients. In all cases except one, however, these ethnicity coefficients are smaller than for the general factor. For example, the Black and Hispanic coefficients for the algebra-specific factor are considerably reduced compared to those for the general factor. An algebra-specific coefficient describes an effect on algebra performance that conditions on the general factor. This type of reduction in coefficients may therefore indicate that the conditioning on the general factor, in addition to conditioning on OTL, to some degree accomplishes the desired conditioning on prior performance.

The Table 11 results for Grade 12 also indicate content-specific effects of OTL variables. As expected, there are particularly strong OTL effects of geometry-trigonometry studies on geometry performance and of algebra-calculus studies on algebra performance. A surprising result is that studying trigonometry has a larger effect on algebra-specific performance than algebra or calculus studies.

Multivariate analysis of process scores in NAEP. Table 12 gives the estimates of the structural model. This model was arrived at as follows. In the measurement part of the model (not reported), the general and specific factors were first allowed to have free loadings. This allowed for a check of the appropriateness of the NAEP classification of items into the problem-solving category. Problem-solving variables with negative loadings on the problem-solving factor were dropped as indicator of this factor and only allowed to load on the general factor. This occurred for 7 of the 29 variables, resulting in the solution shown in Table 12. The variance contribution for the problem-solving factor can be measured as the percentage of the reliable variance that it contributes to problem-solving testlets. The measurement error variance is in this way not involved, which is desirable given that each testlet consists of few items and is therefore quite unreliable. The percentage variance contribution is 63% on average over problem-solving testlets; when conditioning on the background variables, it is 71%. This indicates that the factor is very important.

Table 12 shows that the background variables have about the same effect on the general factor as in Table 10. For the problem-solving factor, student-reported algebra class type has a large positive effect. The problem-solving factor also

BEST COPY AVAILABLE

Table 12

NAEP 1992 Grade 8 Standardized Coefficients (*t*-values) From the Process Analysis (*n*=5013)

	General		Problem solving	
OTL				
(Student-Reported Class Type)				
Prealgebra	0.12	*3.91*	0.05	*1.03*
Algebra	0.59	*13.59*	0.26	*3.86*
(Teacher-Reported)				
NCTM	0.00	*-0.23*	0.02	*0.73*
Class Type				
Enriched	0.26	*8.09*	0.03	*0.56*
Algebra	0.42	*11.12*	-0.03	*-0.47*
Female	-0.01	*-0.51*	-0.04	*-0.85*
Ethnicity				
Black	-0.82	*-17.13*	-0.15	*-2.24*
Hispanic	-0.57	*-11.70*	-0.16	*-2.14*
Asian	0.21	*1.25*	-0.10	*-0.35*
Parents' Ed.	0.15	*10.11*	-0.02	*-1.06*
Type of Community				
Rural	0.06	*1.36*	0.05	*0.58*
Disadv-Urban	-0.29	*-2.61*	0.12	*0.65*
Adv-Urban	0.15	*3.33*	0.23	*3.05*
School Type				
Catholic	0.04	*0.94*	-0.06	*-0.95*
Private	-0.03	*-0.51*	-0.05	*-0.57*
TV Watching	-0.13	*-6.45*	-0.01	*-0.25*
Home Language	0.00	*-0.19*	0.01	*0.18*
R Square	0.41		0.02	

shows differences among Whites, Blacks, and Hispanics, as well as for Type of Community (Advantaged-Urban schools). Similar ethnicity differences were found

in Table 10 for the algebra factor, but a Type of Community effect for specific factors was not found in Table 10. The fact that different effects are found for the problem-solving factor than for the general factor or content-specific factors motivates a further investigation of such achievement components. Related research on reasoning components of math achievement has been carried out using NELS data by Kuppermintz, Ennis, Hamilton, Talbert, and Snow (1994).

Longitudinal analysis of NELS first follow-up data. Table 13 shows that for the general factor at Grade 10, the general factor at Grade 8 is the most important predictor as expected. Despite the inclusion of this prior performance variable, however, student-reported algebra OTL has an additional effect. This effect is present even with reading achievement in Grade 10 taken into account. For the algebra-specific factor in Grade 10, student-reported algebra OTL also has an effect that goes beyond the effect of the prior performance represented by the

Table 13

Standardized Coefficients (*t*-values) From the Longitudinal Model for NELS

	General 8	Algebra 8	General 10	Algebra 10
OTL				
(Student -Reported)				
Algebra (G 10)	—	—	0.23 *9.53*	0.19 *2.20*
(Teacher-Reported)				
Algebra (G 8)	0.45 *9.51*	0.79 *6.38*	0.03 *0.76*	0.07 *0.65*
Performance (G 8)				
General 8	—	—	0.64 *30.46*	0.44 *3.71*
Algebra 8	—	—	0.01 *0.09*	0.16 *1.87*
Reading ability				
Reading 8	0.59 *29.79*	0.20 *3.51*	—	—
Reading 10	—	—	0.32 *21.19*	0.03 *0.64*
R Square	0.44	0.19	0.79	0.33

BEST COPY AVAILABLE

Grade 8 general factor. The partial effects of student-reported algebra OTL are, however, somewhat lower than the corresponding values in Tables 10 and 11 where conditioning on prior performance was not carried out. In part, this may indicate that the conditioning on the general factor carried out in the cross-sectional data analysis of NAEP does not fully accomplish controlling for prior performance as is possible in the longitudinal analysis of NELS. Nevertheless, the conditioning on the general factor is useful also in cross-sectional studies because the general factor is relatively stable over time. In these data, the correlation between the general factor at Grade 8 and Grade 10 is .83. In comparison, we may note that the correlation for the algebra-specific factor at Grade 8 and Grade 10 is only .26.

Discussion

This report suggests a set of analytic approaches that may be used to further the understanding of the relationship between OTL and achievement, and it illustrates them using NAEP and NELS math data. The findings point to three major issues that should be considered in future large-scale educational assessments: OTL sensitivity in items; scoring and reporting of achievement components; and instructions for measuring OTL indicators and relating them to achievement outcomes.

First of all, the concept of OTL-sensitivity in items appears important for future analyses of achievement items where OTL is of interest. In the analysis of the 8th-grade algebra items, it was surprising to find that for a given math achievement level, algebra students did not have a significantly higher probability than nonalgebra students of giving a correct answer to any of the 28 algebra items. The 1992 NAEP Grade 8 algebra items appear not to be sensitive to algebra OTL. This may be a desired effect for a test such as NAEP, which is designed to measure performance in areas of math that are most commonly treated in schools. However, if there is a wish to also be able to use the test as an indicator of effects of curricular change over time, the lack of OTL sensitivity in the items is a deficiency. For example, it may be desirable to add more difficult algebra items or items more specifically geared towards problem solving to capture movements toward better adherence to NCTM standards. If more OTL-sensitive items are added to the test, three important issues arise. First, OTL DIF analysis becomes an important part of test construction to ensure that the items

have the intended characteristics. Second, it becomes necessary to formulate a model for the analysis that incorporates parameters corresponding to the DIF. Third, it becomes important to think of ways to report achievement results covering both OTL-sensitive and OTL-insensitive items.

Second, the analyses point to new possibilities in terms of choice of scoring and reporting of achievement components. The fact that different effects are found for the content-specific factors and for the problem-solving factor than for the general factor motivates further investigations of such achievement components. Interesting analysis possibilities would open up if NAEP "theta" values (latent variable values) could be produced for these additional dimensions; for instance relating algebra-specific skills and problem-solving skills to classroom processes and NCTM reform efforts. Given the sparse matrix sampling of items and an otherwise complex data structure, it is difficult for secondary analysts to prepare reliable performance scores from these various types of items. There is a need for NAEP to provide this. On the other hand, it is not straightforward for NAEP to produce and report such thetas (Mislevy, personal communication, August 1994). For example, the assumption of the IRT model may not be well approximated if content area is ignored, multidimensional IRT modeling for content and process jointly may be cumbersome, and there may be difficulties in reporting that much more detailed information. Perhaps a better place for studies of process dimensions is in surveys such as NELS. It seems worthwhile, however, to consider whether a multidimensional IRT model can be applied either to produce process-related thetas in addition to content-related thetas, or perhaps jointly producing a general theta, content-specific thetas, and process-specific thetas.

Third, it is clear from the previous analyses that it is highly desirable to produce better OTL measures to understand how different opportunities relate to achievement outcomes. These measures need to be both more reliable and more detailed. Multilevel information from students, classrooms, and schools is needed. The careful monitoring of OTL standards and the linking of such information to achievement outcomes make for more useful large-scale educational assessments.

BEST COPY AVAILABLE

References

Baker, D. P. (1993). Compared to Japan, the U.S. is a low achiever...really. New evidence and comment on Westbury. *Educational Researcher*, *22*(3), 18-20.

Bejar, I. I. (1980). A procedure for investigating the unidimensionality of achievement tests based on item parameter estimates. *Journal of Educational Measurement*, *17*, 283-296.

Burstein, L. (1992). *The IEA Study of Mathematics III: Student growth and classroom processes.* Oxford: Pergamon Press.

Burstein, L., & Winters, L. (1994). *Workshop on models for collecting and using opportunity to learn at the state level.* Albuquerque, NM, June.

Holland, P., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson, E., & Carlson, J. (1994). *The NAEP 1992 technical report*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Kifer, E. (1992). Opportunities, talents, and participation. In L. Burstein (Ed.), *The IEA Study of Mathematics III: Student growth and classroom processes* (pp. 279-307). Oxford: Pergamon Press.

Kuppermintz, H., Ennis, M. M., Hamilton, L. S., Talbert, J. E., & Snow, R. E. (1994). *Enhancing the validity and usefulness of large-scale educational assessments: I. NELS:88 Mathematics Achievement* (CRC Tech. Rep.). Stanford, CA: Stanford University.

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, *56*, 177-196.

Mislevy, R. J. (1993). Should "multiple imputations" be treated as "multiple indicators." *Psychometrika*, *58*, 79-85.

Mislevy, R. J., Johnson, E. G., & Muraki, E. (1992). Scaling procedures in NAEP. *Journal of Educational Statistics*, *17*, 131-154.

Muthén, B. (1994). Instructionally sensitive psychometrics: Applications to the Second International Mathematics Study. In I. Westbury, C. A. Ethington, L. A. Sosniak, & D. P. Baker (Eds.), *In search of more effective mathematics education: Examining data from the IEA Second International Mathematics* Study (pp. 293-324). Norwood, NJ: Ablex.

Muthén, B., Kao, C. F., & Burstein, L. (1991). Instructional sensitivity in mathematics achievement item test items: Applications of a new IRT-based detection technique. *Journal of Educational Measurement*, *28*, 1-22.

Muthén, B., Khoo, S-T., & Nelson Goff, G. (1994). *Multidimensional description of subgroup differences in mathematics achievement data from the 1992 National Assessment of Educational Progress* (Tech. Rep.). Los Angeles: University of California, Graduate School of Education & Information Studies.

Silver, E. A., & Kenney P. A. (1993). An examination of relationships between the 1990 NAEP mathematics items for grade 8 and selected themes from the NCTM standards. *Journal for Research in Mathematics Education*, *24*, 159-167.

Westbury, I. (1993). American and Japanese achievement...again. A response to Baker. *Educational Researcher*, *22*(3), 21-25.

Appendix

Input for Structural Equation Modeling Using the LISCOMP Computer Program

```
TI NAEP 92 GRADE 8 TABLE 10 IMPUTATION 1
DA IY=22 NO=4715
MO MO=SE P3 NE=23 LY=FI TE=FI PS=FI BE=FI
LL
  'GENERAL' 'NUMOP' 'MEAS' 'GEOM' 'DATA' 'ALGEB' 'FEMALE' 'BLACK'
  'HISP' 'ASIAN' 'PARED' 'RURAL' 'DISADV' 'ADV' 'CATH' 'PRIV' 'TV'
  'LANGHOM' 'PREALG' 'ALGEB' 'FNCTM' 'ENRICH' 'TALG'

VA 1.0 LY(2,1)                                                LY(2,3)   LY(3,4)
       LY(4,5)   LY( 5,6)  LY( 6,7)  LY( 7,8)  LY( 8,9)  LY( 9,10) LY(10,11)
       LY(11,12) LY(12,13) LY(13,14) LY(14,15) LY(15,16) LY(16,17)
       LY(17,18) LY(18,19) LY(19,20) LY(20,21) LY(21,22) LY(22,23)

FR     LY(1,1) LY(3,1) LY(4,1) LY(5,1)
VA 0.8 LY(1,1) LY(3,1) LY(4,1) LY(5,1)

FR BE(1,7)  BE(1,8)  BE(1,9)  BE(1,10) BE(1,11) BE(1,12) BE(1,13) BE(1,14)
   BE(1,15) BE(1,16) BE(1,17) BE(1,18) BE(1,19) BE(1,20) BE(1,21)
   BE(1,22) BE(1,23)
FR BE(3,7)  BE(3,8)  BE(3,9)  BE(3,10) BE(3,11) BE(3,12) BE(3,13) BE(3,14)
   BE(3,15) BE(3,16) BE(3,17) BE(3,18) BE(3,19) BE(3,20) BE(3,21)
   BE(3,22) BE(3,23)
FR BE(4,7)  BE(4,8)  BE(4,9)  BE(4,10) BE(4,11) BE(4,12) BE(4,13) BE(4,14)
   BE(4,15) BE(4,16) BE(4,17) BE(4,18) BE(4,19) BE(4,20) BE(4,21)
   BE(4,22) BE(4,23)
FR BE(5,7)  BE(5,8)  BE(5,9)  BE(5,10) BE(5,11) BE(5,12) BE(5,13) BE(5,14)
   BE(5,15) BE(5,16) BE(5,17) BE(5,18) BE(5,19) BE(5,20) BE(5,21)
   BE(5,22) BE(5,23)
FR BE(6,7)  BE(6,8)  BE(6,9)  BE(6,10) BE(6,11) BE(6,12) BE(6,13) BE(6,14)
   BE(6,15) BE(6,16) BE(6,17) BE(6,18) BE(6,19) BE(6,20) BE(6,21)
   BE(6,22) BE(6,23)
FR PS(1,1)          PS(3,3) PS(4,4) PS(5,5) PS(6,6)

VA 0.5 PS(1,1)          PS(3,3) PS(4,4) PS(5,5) PS(6,6)

VA  .250 PS(7,7)
VA  .000 PS(8,7)
VA  .001 PS(9,7)
VA -.001 PS(10,7)
VA -.035 PS(11,7)
VA  .000 PS(12,7)
VA -.002 PS(13,7)
VA -.003 PS(14,7)
VA  .004 PS(15,7)
VA -.006 PS(16,7)
VA -.015 PS(17,7)
VA  .007 PS(18,7)
VA  .001 PS(19,7)
VA -.001 PS(20,7)
VA  .005 PS(21,7)
VA  .000 PS(22,7)
VA  .000 PS(23,7)
VA  .133 PS(8,8)
VA -.016 PS( 9,8)
```

```
VA -.005 PS(10,8)
VA -.009 PS(11,8)
VA -.003 PS(12,8)
VA  .017 PS(13,8)
VA -.008 PS(14,8)
VA -.004 PS(15,8)
VA -.010 PS(16,8)
VA  .070 PS(17,8)
VA -.012 PS(18,8)
VA -.001 PS(19,8)
VA -.014 PS(20,8)
VA  .023 PS(21,8)
VA  .011 PS(22,8)
VA -.019 PS(23,8)
VA  .092 PS( 9,9)
VA -.003 PS(10,9)
VA -.068 PS(11,9)
VA -.004 PS(12,9)
VA  .014 PS(13,9)
VA -.009 PS(14,9)
VA -.005 PS(15,9)
VA -.006 PS(16,9)
VA  .016 PS(17,9)
VA  .090 PS(18,9)
VA -.004 PS(19,9)
VA -.007 PS(20,9)
VA  .020 PS(21,9)
VA  .015 PS(22,9)
VA -.010 PS(23,9)
VA  .028 PS(10,10)
VA  .016 PS(11,10)
VA -.003 PS(12,10)
VA  .001 PS(13,10)
VA  .004 PS(14,10)
VA  .001 PS(15,10)
VA  .002 PS(16,10)
VA -.006 PS(17,10)
VA  .026 PS(18,10)
VA -.003 PS(19,10)
VA  .009 PS(20,10)
VA  .005 PS(21,10)
VA -.001 PS(22,10)
VA  .007 PS(23,10)
VA 1.041 PS(11,11)
VA -.028 PS(12,11)
VA -.037 PS(13,11)
VA  .055 PS(14,11)
VA  .032 PS(15,11)
VA  .037 PS(16,11)
VA -.096 PS(17,11)
VA -.031 PS(18,11)
VA  .010 PS(19,11)
VA  .091 PS(20,11)
VA -.001 PS(21,11)
VA -.032 PS(22,11)
VA  .106 PS(23,11)
VA  .080 PS(12,12)
VA -.007 PS(13,12)
VA -.009 PS(14,12)
VA -.006 PS(15,12)
VA -.007 PS(16,12)
VA  .003 PS(17,12)
VA -.014 PS(18,12)
VA  .010 PS(19,12)
```

BEST COPY AVAILABLE

```
VA -.010 PS(20,12)
VA  .015 PS(21,12)
VA -.016 PS(22,12)
VA -.006 PS(23,12)
VA  .073 PS(13,13)
VA -.008 PS(14,13)
VA -.005 PS(15,13)
VA -.004 PS(16,13)
VA  .025 PS(17,13)
VA  .025 PS(18,13)
VA -.011 PS(19,13)
VA -.002 PS(20,13)
VA  .000 PS(21,13)
VA  .007 PS(22,13)
VA -.011 PS(23,13)
VA  .091 PS(14,14)
VA  .005 PS(15,14)
VA  .028 PS(16,14)
VA -.030 PS(17,14)
VA  .006 PS(18,14)
VA -.001 PS(19,14)
VA  .018 PS(20,14)
VA -.004 PS(21,14)
VA -.011 PS(22,14)
VA  .025 PS(23,14)
VA  .110 PS(15,15)
VA -.010 PS(16,15)
VA -.009 PS(17,15)
VA -.003 PS(18,15)
VA -.003 PS(19,15)
VA  .001 PS(20,15)
VA -.003 PS(21,15)
VA  .004 PS(22,15)
VA  .012 PS(23,15)
VA  .073 PS(16,16)
VA -.025 PS(17,16)
VA  .003 PS(18,16)
VA  .004 PS(19,16)
VA  .006 PS(20,16)
VA -.015 PS(21,16)
VA -.008 PS(22,16)
VA  .012 PS(23,16)
VA  .450 PS(17,17)
VA  .010 PS(18,17)
VA  .006 PS(19,17)
VA -.043 PS(20,17)
VA  .000 PS(21,17)
VA  .016 PS(22,17)
VA -.037 PS(23,17)
VA  .428 PS(18,18)
VA -.014 PS(19,18)
VA  .013 PS(20,18)
VA  .032 PS(21,18)
VA  .019 PS(22,18)
VA -.002 PS(23,18)
VA  .197 PS(19,19)
VA -.062 PS(20,19)
VA -.012 PS(21,19)
VA -.012 PS(22,19)
VA  .000 PS(23,19)
VA  .177 PS(20,20)
VA  .050 PS(21,20)
VA -.044 PS(22,20)
VA  .100 PS(23,20)
```

```
VA 1.000 PS(21,21)
VA -.018 PS(22,21)
VA  .035 PS(23,21)
VA  .211 PS(22,22)
VA -.100 PS(23,22)
VA  .221 PS(23,23)

OU MN ES SE TV ET VE RS
RA FO UN=8
(F5.2,T26,F5.2,T51,F5.2,T76,F5.2,T101,F5.2/14F1.0,F10.3,2F1.0)
```